

University of Groningen

## The visibility-based tapered gridded estimator (TGE) for the redshifted 21-cm power spectrum

Choudhuri, Samir; Bharadwaj, Somnath; Chatterjee, Suman; Ali, Sk. Saiyad; Roy, Nirupam; Ghosh, Abhik

*Published in:*  
Monthly Notices of the Royal Astronomical Society

*DOI:*  
[10.1093/mnras/stw2254](https://doi.org/10.1093/mnras/stw2254)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2016

[Link to publication in University of Groningen/UMCG research database](#)

### *Citation for published version (APA):*

Choudhuri, S., Bharadwaj, S., Chatterjee, S., Ali, S. S., Roy, N., & Ghosh, A. (2016). The visibility-based tapered gridded estimator (TGE) for the redshifted 21-cm power spectrum. *Monthly Notices of the Royal Astronomical Society*, 463(4), 4093-4107. <https://doi.org/10.1093/mnras/stw2254>

### **Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### **Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*



# The visibility-based tapered gridded estimator (TGE) for the redshifted 21-cm power spectrum

Samir Choudhuri,<sup>1</sup>★ Somnath Bharadwaj,<sup>1</sup> Suman Chatterjee,<sup>1</sup> Sk. Saiyad Ali,<sup>2</sup> Nirupam Roy<sup>3</sup> and Abhik Ghosh<sup>4</sup>

<sup>1</sup>Department of Physics, and Centre for Theoretical Studies, IIT Kharagpur, PIN 721 302, India

<sup>2</sup>Department of Physics, Jadavpur University, Kolkata 700032, India

<sup>3</sup>Department of Physics, Indian Institute of Science, Bangalore 560012, India

<sup>4</sup>Kapteyn Astronomical Institute, PO Box 800, NL-9700 AV Groningen, the Netherlands

Accepted 2016 September 5. Received 2016 September 5; in original form 2016 June 28

## ABSTRACT

We present an improved visibility-based tapered gridded estimator (TGE) for the power spectrum of the diffuse sky signal. The visibilities are gridded to reduce the total computation time for the calculation, and tapered through a convolution to suppress the contribution from the outer regions of the telescope's field of view. The TGE also internally estimates the noise bias, and subtracts this out to give an unbiased estimate of the power spectrum. An earlier version of the 2D TGE for the angular power spectrum  $C_\ell$  is improved and then extended to obtain the 3D TGE for the power spectrum  $P(\mathbf{k})$  of the 21-cm brightness temperature fluctuations. Analytic formulas are also presented for predicting the variance of the binned power spectrum. The estimator and its variance predictions are validated using simulations of 150-MHz Giant Metrewave Radio Telescope (GMRT) observations. We find that the estimator accurately recovers the input model for the 1D spherical power spectrum  $P(k)$  and the 2D cylindrical power spectrum  $P(k_\perp, k_\parallel)$ , and that the predicted variance is in reasonably good agreement with the simulations.

**Key words:** methods: data analysis – methods: statistical – techniques: interferometric – diffuse radiation.

## 1 INTRODUCTION

Observations of the redshifted neutral hydrogen (H I) 21-cm radiation have the potential to probe a wide range of cosmological and astrophysical phenomena over a large redshift range,  $0 < z \lesssim 200$  (Bharadwaj & Ali 2005; Furlanetto, Oh & Briggs 2006; Morales & Wyithe 2010; Prichard & Loeb 2012; Mellema et al. 2013). There now are several ongoing experiments, such as the Donald C. Backer Precision Array to Probe the Epoch of Reionization (PAPER;<sup>1</sup> Parsons et al. 2010), the Low Frequency Array (LOFAR;<sup>2</sup> van Haarlem et al. 2013; Yatawatta et al. 2013) and the Murchison Wide-field Array (MWA;<sup>3</sup> Bowman et al. 2013; Tingay et al. 2013), that aim to measure the power spectrum of the 21-cm radiation from the Epoch of Reionization (EoR,  $6 \lesssim z \lesssim 13$ ). Future telescopes such as the Square Kilometre Array (SKA1 LOW;<sup>4</sup> Koopmans et al. 2015) and the Hydrogen Epoch of

Reionization Array (HERA;<sup>5</sup> Neben et al. 2016) are planned in order to achieve even higher sensitivity for measuring the EoR 21-cm power spectrum. Several other upcoming experiments such as the Ooty Wide Field Array (OWFA; Prasad & Subrahmanya 2011; Ali & Bharadwaj 2014), the Canadian Hydrogen Intensity Mapping Experiment (CHIME<sup>6</sup>; Bandura et al. 2014), the Baryon Acoustic Oscillation Broad-band, Broad Beam Array (BAOBAB;<sup>7</sup> Pober et al. 2013a) and the Square Kilometre Array (SKA1 MID; Bull et al. 2015) target the post-reionization 21-cm signal ( $0 < z \lesssim 6$ ).

In spite of the availability of sensitive new instruments, the main challenge still arises from the fact that the cosmological 21-cm signal is buried in astrophysical foregrounds that are 4–5 orders of magnitude brighter (Shaver et al. 1999; Di Matteo et al. 2002; Santos, Cooray & Knox 2005; Ali, Bharadwaj & Chengalur 2008; Paciga et al. 2011; Ghosh et al. 2011a,b). A large variety of techniques have been proposed to overcome this problem and estimate the 21-cm power spectrum. The various approaches may be broadly

\* E-mail: samir11@phy.iitkgp.ernet.in

<sup>1</sup> <http://astro.berkeley.edu/dbacker/eor>

<sup>2</sup> <http://www.lofar.org/>

<sup>3</sup> <http://www.mwatelescope.org>

<sup>4</sup> <http://www.skatelescope.org/>

<sup>5</sup> <http://reionization.org/>

<sup>6</sup> <http://chime.phas.ubc.ca/>

<sup>7</sup> <http://bao.berkeley.edu/>

divided into two classes, namely (1) foreground removal, and (2) foreground avoidance.

The idea in foreground removal is to model the foregrounds and subtract these out either directly from the data (e.g. Ali et al. 2008) or from the power spectrum estimator after correlating the data (e.g. Ghosh et al. 2011a,b). Foreground removal is a topic of intense current research (Jelić et al. 2008; Bowman, Morales & Hewitt 2009; Paciga et al. 2011; Chapman et al. 2012; Parsons et al. 2012; Liu & Tegmark 2012; Trott, Wayth & Tingay 2012; Pober et al. 2013b; Paciga et al. 2013; Parsons et al. 2014; Trott et al. 2016).

Various studies (e.g. Datta, Bowman & Carilli 2010) have shown that the foreground contribution to the cylindrical power spectrum  $P(k_{\perp}, k_{\parallel})$  is expected to be restricted within a wedge in the 2D  $(k_{\perp}, k_{\parallel})$ -plane. The idea in foreground avoidance is to avoid the Fourier modes within the foreground wedge and use only the uncontaminated modes outside the wedge to estimate the 21-cm power spectrum (Vedantham, Udaya Shankar & Subrahmanyam 2012; Thyagarajan et al. 2013; Dillon et al. 2014, 2015; Liu, Parsons & Trott 2014a,b; Pober et al. 2014; Ali et al. 2015). In a recent paper, Jacobs et al. (2016) compared several power spectrum estimation techniques in the context of the MWA.

Point sources dominate the low-frequency sky at the angular scales  $\leq 4^\circ$  (Ali et al. 2008) that are relevant for the EoR 21-cm power spectrum with telescopes such as the GMRT, LOFAR and the upcoming SKA. It is difficult to model and subtract the point sources that are located at the periphery of the telescope's field of view (FoV). The antenna response deviates from circular symmetry, and is highly frequency- and time-dependent in the outer regions of the telescope's FoV. The calibration also differs from the phase centre owing to ionospheric fluctuations. The residual point sources located far from the phase centre cause the signal to oscillate along the frequency direction (Ghosh et al. 2011a,b). This poses a severe problem for foreground removal techniques, which assume a smooth behaviour of the signal along the frequency direction. Equivalently, these distant point sources reduce the EoR window by increasing the area under the foreground wedge in  $(k_{\perp}, k_{\parallel})$ -space (Thyagarajan et al. 2015). In a recent paper, Pober et al. (2016) showed that it is important to correctly model and subtract the distant point sources in order to detect the redshifted 21-cm signal. Point source subtraction is also important for measuring the angular power spectrum of the diffuse Galactic synchrotron radiation (Bernardi et al. 2009; Ghosh et al. 2012; Iacubelli et al. 2013). Apart from being an important foreground component for the EoR 21-cm signal, this is also interesting in its own right.

It is possible to suppress the contribution from the outer parts of the telescope's FoV by tapering the sky response through a suitably chosen window function. Ghosh et al. (2011b) analysed 610-MHz GMRT data to show that it is possible to implement the tapering by convolving the observed visibilities with the Fourier transform of the window function. It is found that this reduces the amplitude of the oscillation along the frequency direction. Our earlier work, Choudhuri et al. (2014, hereafter Paper I), introduced the tapered gridded estimator (TGE), which places the findings of Ghosh et al. (2011b) on a sound theoretical footing. Considering observations at a single frequency, the TGE estimates the angular power spectrum  $C_{\ell}$  of the 2D sky signal directly from the measured visibilities while simultaneously tapering the sky response. As a test-bed for the TGE, Paper I considered a situation in which the point sources were identified and subtracted out so that the residual visibilities are dominated by the Galactic synchrotron radiation. This has been used to investigate how well the TGE is able to recover the angular power spectrum of the input model used to simulate the Galactic

synchrotron emission at 150 MHz. While most of the analysis was for the GMRT, simulations for LOFAR were also considered. These investigations showed that the TGE is able to recover the input model  $C_{\ell}^M$  to a high level of precision, provided that the baselines have a uniform  $uv$  coverage. For the GMRT, which has a patchy  $uv$  coverage, the  $C_{\ell}$  is slightly overestimated using TGE, although the excess is largely within the  $1\sigma$  errors. This deviation is found to be reduced in cases with a more uniform and denser baseline distribution, such as LOFAR. Paper I also analysed the effects of gain errors and the  $w$ -term.

In a recent paper, Choudhuri et al. (2016, hereafter Paper II), we further developed the simulations of Paper I to include point sources. We used conventional radio astronomical techniques to model and subtract the point sources from the central region of the primary beam. As detailed in Paper II, it is difficult to do this for sources that are far from the phase centre, and these persist as residuals in the visibility data. We found that these residual point sources dominate the  $C_{\ell}$  estimated at large baselines. We also showed that it is possible to suppress the contribution from these residual sources located at the periphery of the FoV by using TGE with a suitably chosen window function.

Removing the noise bias is an important issue for any power spectrum estimator. As demonstrated in Paper II, the TGE internally estimates the actual noise bias from the data and subtracts this out to give an unbiased estimate of the power spectrum.

In the present work we report progress in two areas. First, our earlier implementation of the TGE assumed a uniform and dense baseline  $uv$  coverage in calculations of the normalization coefficient, which relates visibility correlations to the estimated angular power spectrum  $C_{\ell}$ . We found (Paper I), however, that this leads to an overestimate of  $C_{\ell}$  for instruments such as the GMRT that have a sparse and patchy  $uv$  coverage. In Section 2 of this paper we present an improved TGE that overcomes this problem by using simulations to estimate the normalization coefficient. Second, the entire analysis of Papers I and II was restricted to observations at a single frequency, wherein the pertinent problem was to quantify the 2D angular fluctuations of the sky signal. This approach, however, is inadequate for the 3D redshifted H I 21-cm signal, for which it is necessary to simultaneously quantify the fluctuations along the frequency direction. In Section 3 of this paper we generalize the TGE in order to quantify the 3D 21-cm signal and estimate the spatial power spectrum of the 21-cm brightness temperature fluctuations  $P(\mathbf{k})$ . We discuss two binning schemes, which respectively yield the spherically averaged (1D) power spectrum  $P(k)$  and the cylindrically averaged (2D) power spectrum  $P(k_{\perp}, k_{\parallel})$ , and present theoretical expressions for predicting the expected variance. We validated the estimator and its variance predictions using simulations, which are described in Section 4 and for which the results are presented in Section 5. Section 6 presents the summary and conclusions.

In this paper, we have used cosmological parameters from the (Planck + WMAP) best-fit  $\Lambda$ CDM cosmology (Planck Collaboration XIII 2016).

## 2 $C_{\ell}$ ESTIMATION

### 2.1 An improved TGE

In this section we restrict our attention to a single frequency channel  $\nu_a$ , which we do not show explicitly in any of the subsequent equations. The measured visibilities  $\mathcal{V}_i$  can be decomposed into two contributions, namely

$$\mathcal{V}_i = \mathcal{S}(\mathbf{U}_i) + \mathcal{N}_i, \quad (1)$$

the sky signal and system noise respectively, where  $\mathbf{U}_i$  is the baseline corresponding to the  $i$ th visibility. The signal contribution  $\mathcal{S}(\mathbf{U}_i)$  records the Fourier transform of the product of the telescope's primary beam pattern  $\mathcal{A}(\theta)$  and the specific intensity fluctuation on the sky  $\delta I(\theta)$ . Expressing the signal in terms of brightness temperature fluctuations  $\delta T(\theta)$  we have

$$\mathcal{S}(\mathbf{U}_i) = \left( \frac{\partial B}{\partial T} \right) \int d^2\theta e^{2\pi i \mathbf{U}_i \cdot \theta} \mathcal{A}(\theta) \delta T(\theta), \quad (2)$$

where  $B = 2k_B T / \lambda^2$  is the Planck function in the Rayleigh–Jeans limit, which is valid at the frequencies of interest here. In terms of Fourier components we have

$$\mathcal{S}(\mathbf{U}_i) = \left( \frac{\partial B}{\partial T} \right) \int d^2U \tilde{a}(\mathbf{U}_i - \mathbf{U}) \Delta \tilde{T}(\mathbf{U}), \quad (3)$$

where  $\Delta \tilde{T}(\mathbf{U})$  and  $\tilde{a}(\mathbf{U})$  are the Fourier transforms of  $\delta T(\theta)$  and  $\mathcal{A}(\theta)$  respectively. Here we assume that  $\delta T(\theta)$  is a particular realization of a statistically homogeneous and isotropic Gaussian random process on the sky. Its statistical properties are completely characterized by the angular power spectrum of the brightness temperature fluctuations  $C_\ell$  defined through

$$\langle \Delta \tilde{T}(\mathbf{U}) \Delta \tilde{T}^*(\mathbf{U}') \rangle = \delta_D^2(\mathbf{U} - \mathbf{U}') C_{2\pi U}, \quad (4)$$

where  $\delta_D^2(\mathbf{U} - \mathbf{U}')$  is a 2D Dirac delta function and  $2\pi U = \ell$  is the angular multipole. The angular brackets  $\langle \dots \rangle$  here denote an ensemble average over different realizations of the stochastic temperature fluctuations on the sky.

The noise in the different visibilities is uncorrelated, and we have

$$\langle \mathcal{V}_i \mathcal{V}_j \rangle = \langle \mathcal{S}_i \mathcal{S}_j \rangle + \langle |\mathcal{N}_i|^2 \rangle \delta_{i,j}, \quad (5)$$

where  $\langle |\mathcal{N}_i|^2 \rangle$  is the noise variance of the visibilities,  $\delta_{i,j}$  is a Kronecker delta, and

$$\langle \mathcal{S}_i \mathcal{S}_j \rangle = \left( \frac{\partial B}{\partial T} \right)^2 \int d^2U \tilde{a}(\mathbf{U}_i - \mathbf{U}) \tilde{a}^*(\mathbf{U}_j - \mathbf{U}) C_{2\pi U_i}. \quad (6)$$

This convolution can be approximated by a multiplying factor if  $C_{2\pi U}$  is nearly constant across the width of  $\tilde{a}(\mathbf{U}_i - \mathbf{U})$ , which is the situation for large baselines where the antenna separation is large compared with the telescope diameter (Paper I), and we have

$$\langle |\mathcal{V}_i|^2 \rangle = V_0 C_{2\pi U_i} + \langle |\mathcal{N}_i|^2 \rangle, \quad (7)$$

where

$$V_0 = \left( \frac{\partial B}{\partial T} \right)^2 \int d^2U |\tilde{a}(\mathbf{U}_i - \mathbf{U})|^2. \quad (8)$$

We see that the correlation of a visibility with itself provides an estimate of the angular power spectrum, except for the terms  $\langle |\mathcal{N}_i|^2 \rangle$ , which introduce a positive noise bias.

It is possible to control the sidelobe response of the telescope's beam pattern  $\mathcal{A}(\theta)$  by tapering the sky response through a frequency-independent window function  $\mathcal{W}(\theta)$ . In this work we use a Gaussian  $\mathcal{W}(\theta) = \exp(-\theta^2/\theta_w^2)$ , with  $\theta_w$  chosen so that the window function cuts off the sky response well before the first null of  $\mathcal{A}(\theta)$ . This tapering is achieved by convolving the measured visibilities with the Fourier transform of  $\mathcal{W}(\theta)$ . We choose a rectangular grid in the  $uv$ -plane and consider the convolved visibilities

$$\mathcal{V}_{cg} = \sum_i \tilde{w}(\mathbf{U}_g - \mathbf{U}_i) \mathcal{V}_i, \quad (9)$$

where  $\tilde{w}(\mathbf{U}) = \pi \theta_w^2 \exp(-\pi^2 U^2 \theta_w^2)$  is the Fourier transform of  $\mathcal{W}(\theta)$  and  $\mathbf{U}_g$  refers to the different grid points. As shown in Paper I,

gridding reduces the computation in comparison with an estimator that uses pairs of visibilities to estimate the power spectrum. We now focus our attention on  $\mathcal{S}_{cg}$ , which is the sky signal contribution to  $\mathcal{V}_{cg}$ . This can be written as

$$\mathcal{S}_{cg} = \left( \frac{\partial B}{\partial T} \right) \int d^2U \tilde{K}(\mathbf{U}_g - \mathbf{U}) \Delta \tilde{T}(\mathbf{U}), \quad (10)$$

where

$$\tilde{K}(\mathbf{U}_g - \mathbf{U}) = \int d^2U' \tilde{w}(\mathbf{U}_g - \mathbf{U}') B(\mathbf{U}') \tilde{a}(\mathbf{U}' - \mathbf{U}) \quad (11)$$

is an effective 'gridding kernel', and

$$B(\mathbf{U}) = \sum_i \delta_D^2(\mathbf{U} - \mathbf{U}_i) \quad (12)$$

is the baseline sampling function of the measured visibilities.

Proceeding in exactly the same way as we did for equation (7), we have

$$\langle |\mathcal{V}_{cg}|^2 \rangle = V_{1g} C_{2\pi U_g} + \sum_i |\tilde{w}(\mathbf{U}_g - \mathbf{U}_i)|^2 \langle |\mathcal{N}_i|^2 \rangle, \quad (13)$$

where

$$V_{1g} = \left( \frac{\partial B}{\partial T} \right)^2 \int d^2U |\tilde{K}(\mathbf{U}_i - \mathbf{U})|^2. \quad (14)$$

Here again we see that the correlation of the tapered gridded visibility with itself provides an estimate of the angular power spectrum, except for the terms  $\langle |\mathcal{N}_i|^2 \rangle$ , which introduce a positive noise bias.

Combining equations (7) and (13), we have

$$\left\langle \left( |\mathcal{V}_{cg}|^2 - \sum_i |\tilde{w}(\mathbf{U}_g - \mathbf{U}_i)|^2 |\mathcal{V}_i|^2 \right) \right\rangle = M_g C_{2\pi U_g}, \quad (15)$$

where

$$M_g = V_{1g} - \sum_i |\tilde{w}(\mathbf{U}_g - \mathbf{U}_i)|^2 V_0. \quad (16)$$

This allows us to define the tapered gridded estimator (TGE) as

$$\hat{E}_g = M_g^{-1} \left( |\mathcal{V}_{cg}|^2 - \sum_i |\tilde{w}(\mathbf{U}_g - \mathbf{U}_i)|^2 |\mathcal{V}_i|^2 \right). \quad (17)$$

The TGE defined here (equation 17) incorporates three novel features, as follows. First, the estimator uses the gridded visibilities to estimate  $C_\ell$ ; this is computationally much faster than individually correlating the visibilities. Second, the correlation of the gridded visibilities is used to estimate  $C_\ell$ . A positive noise bias is removed by subtracting the auto-correlation of the visibilities. Third, the estimator allows us to taper the FoV so as to restrict the contribution from the sources in the outer regions and the sidelobes. It is, however, necessary to note that this comes at a cost, which we now discuss. First, we lose information at the largest angular scales owing to the reduced FoV. This restricts the smallest  $\ell$  value at which it is possible to estimate the power spectrum. Second, the reduced FoV results in a larger cosmic variance for the smaller angular modes that are within the tapered FoV.

The TGE provides an unbiased estimate of  $C_{\ell_g}$  at the angular multipole  $\ell_g = 2\pi U_g$ ; that is,

$$\langle \hat{E}_g \rangle = C_{\ell_g}. \quad (18)$$

We use this to define the binned TGE for bin  $a$  as

$$\hat{E}_G(a) = \frac{\sum_g w_g \hat{E}_g}{\sum_g w_g}, \quad (19)$$

where  $w_g$  refers to the weight assigned to the contribution from any particular grid point. In the entire subsequent analysis we use the weight  $w_g = 1$ , which assigns equal weight to all the grid points that are sampled by the baselines.

The binned estimator has an expectation value

$$\bar{C}_{\ell_a} = \frac{\sum_g w_g C_{\ell_g}}{\sum_g w_g}, \quad (20)$$

where  $\bar{C}_{\ell_a}$  is the average angular power spectrum at

$$\bar{\ell}_a = \frac{\sum_g w_g \ell_g}{\sum_g w_g}, \quad (21)$$

which is the effective angular multipole for bin  $a$ .

## 2.2 Calculating $M_g$

The discussion thus far has not addressed how to calculate  $M_g$ , which is the normalization constant for the TGE (equation 17). The values of  $M_g$  (equation 16) depend on the baseline distribution (equation 12) and the form of the tapering function  $\mathcal{W}(\theta)$ , and it is necessary to calculate  $M_g$  at every grid point in the  $uv$ -plane. Our earlier work (Paper I) presented an analytic approximation with which it is possible to estimate  $M_g$ . While this was found to work very well in situations in which the baselines have a nearly uniform and dense  $uv$  coverage (fig. 7 of Paper I), it leads to an overestimate of  $C_\ell$  if there is a sparse and non-uniform  $uv$  coverage. Here we present a different method to estimate  $M_g$ , which, as we show later, works very well even with a sparse and non-uniform  $uv$  coverage.

We proceed by calculating simulated visibilities  $[\mathcal{V}_i]_{\text{UPAS}}$  corresponding to a unit angular power spectrum (UAPS) that has  $C_\ell = 1$  with exactly the same baseline distribution as the actual observed visibilities. We then have (equation 15)

$$M_g = \left\langle \left( |\mathcal{V}_{cg}|^2 - \sum_i |\tilde{w}(\mathbf{U}_g - \mathbf{U}_i)|^2 (|\mathcal{V}_i|^2) \right) \right\rangle_{\text{UPAS}}, \quad (22)$$

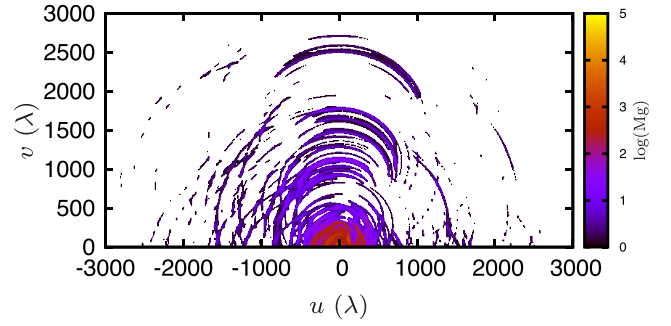
which allows us to estimate  $M_g$ . We average over  $N_u$  independent realizations of the UPAS in order to reduce the statistical uncertainty ( $\delta M_g / M_g \sim 1/\sqrt{N_u}$ ) in the estimated  $M_g$ .

## 2.3 Validating the estimator

We tested the entire method of analysis using simulations of 8 h of 150-MHz GMRT observations targeted on an arbitrarily selected field located at RA =  $10^{\text{h}} 46^{\text{m}} 00^{\text{s}}$  and Dec. =  $59^\circ 00' 59''$ . The simulations only incorporate the diffuse Galactic synchrotron radiation, for which we use the measured angular power spectrum (Ghosh et al. 2012)

$$C_\ell^{\text{M}} = A_{150} \left( \frac{1000}{\ell} \right)^\beta \quad (23)$$

as the input model to generate the brightness temperature fluctuations on the sky. Here  $A_{150} = 513 \text{ mK}^2$  and  $\beta = 2.34$  (Ghosh et al. 2012). The simulation covers a region of the sky of size  $\sim 26.4 \times 26.4$ , which is slightly more than 10 times the full width at half-maximum (FWHM) of the GMRT primary beam ( $\theta_{\text{FWHM}} = 157'$ ). The diffuse signal was simulated on a grid of resolution  $\sim 0.5'$ , and the entire analysis was restricted to baselines within  $U \leq 3000$ . Our earlier work (Paper II) and also the discussion in this paper show that the noise bias cancels out from the TGE, and we have not included the system noise in these simulations.



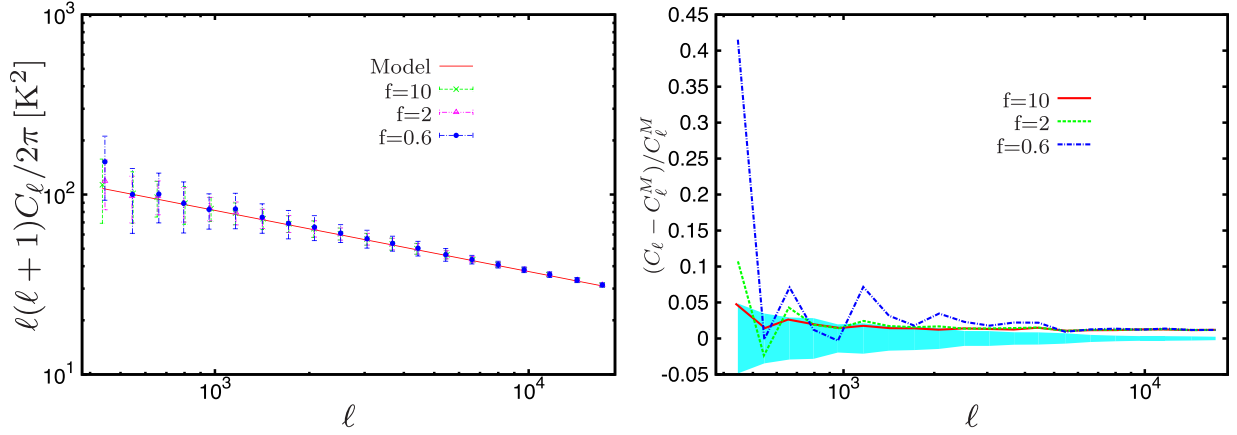
**Figure 1.** Plot of  $M_g$  for a fixed value of  $f = 0.6$ . Note that the baselines in the lower half of the  $uv$ -plane have been folded onto the upper half.

We modelled the tapering window function as a Gaussian  $\mathcal{W}(\theta) = \exp(-\theta^2/\theta_w^2)$ , where we parametrize  $\theta_w = f\theta_0$ , where  $\theta_0 = 0.6 \times \theta_{\text{FWHM}}$ , and preferably  $f \leq 1$  so that  $\mathcal{W}(\theta)$  cuts off the sky response well before the first null of the primary beam. After tapering, we have an effective beam pattern  $\mathcal{A}_w(\theta) = \mathcal{W}(\theta) \mathcal{A}(\theta, \nu)$ , which is well approximated by a Gaussian,  $\mathcal{A}_{\text{vw}}(\theta) = \exp(-\theta^2/\theta_1^2)$  with  $\theta_1 = f(1 + f^2)^{-1/2}\theta_0$ . The spacing of the  $uv$ -grid required for the TGE is decided by  $\tilde{a}_w(U) = \pi\theta_1^2 \exp(-\pi^2 U^2 \theta_1^2)$ , which is the Fourier transform of  $\mathcal{A}_w(\theta)$ . We have chosen a grid spacing  $\Delta U = \sqrt{\ln 2}/(2\pi\theta_1)$ , which corresponds to one-quarter of the FWHM of  $\tilde{a}_w(U)$ . The convolution in equation (9) was restricted to the visibilities within a disc of radius  $12 \Delta U$  around each grid point. The function  $\tilde{w}(\mathbf{U}_g - \mathbf{U}_i)$  falls off rapidly, and we do not expect the visibilities beyond this to make a significant contribution.

We considered three values,  $f = 10, 2$  and  $0.6$ , for the tapering, where  $f = 10$  essentially corresponds to a situation with no tapering, and the sky response is confined to a progressively smaller region as the value of  $f$  is reduced to  $f = 2.0$  and  $0.6$  respectively (see fig. 1 of Paper II). We used  $N_u = 128$  independent realizations of the UAPS to estimate  $M_g$  for each point in the  $uv$ -grid. It is necessary to calculate  $M_g$  for each value of  $f$  separately. Fig. 1 shows the values of  $M_g$  for  $f = 0.6$ . It can be seen that this roughly traces out the  $uv$ -tracks of the baselines; the convolution with  $\tilde{w}(\mathbf{U}_g - \mathbf{U}_i)$  results in a thickening of the tracks. The values of  $M_g$  are roughly proportional to  $N_g^2 - N_g$ , where  $N_g$  is the number of visibilities that contribute to any particular grid point.

The estimator (equation 17) was applied to the simulated visibility data that was generated using the model angular power spectrum (equation 23). The estimated angular power spectrum was binned into 20 annular bins of equal logarithmic spacing. We used  $N_r = 128$  independent realizations of the simulation to calculate the mean and standard deviation of  $C_\ell$  shown in the left panel of Fig. 2. It can be seen that the TGE is able to recover the input model  $C_\ell^{\text{M}}$  quite accurately. As noted earlier, our previous implementation of TGE (Paper I) had a problem in that the estimated  $C_\ell$  was in all cases in excess of the input model  $C_\ell^{\text{M}}$ , although the deviations were within the  $1\sigma$  error bars throughout. The right panel of Fig. 2 shows the fractional deviation  $(C_\ell - C_\ell^{\text{M}})/C_\ell^{\text{M}}$  for the improved TGE introduced in this paper for the three values of  $f$  mentioned above. It can be seen that for all the values of  $f$ , the fractional deviation is less than 10 per cent for  $\ell \geq 500$ . This is a considerable improvement on the results of Paper I, where we had deviations of 20 to 50 per cent. The fractional deviation is seen to increase as we increase the tapering, that is, reduce the value of  $f$ . It can be seen that for  $f = 10$  and  $2$ , the fractional deviation is less than 3 per cent for all values of  $\ell$  except at the smallest bin. The fractional deviation for  $f = 0.6$  is less than 5 per cent except at the smallest value of  $\ell$ , where it becomes





**Figure 2.** The left panel shows a comparison of the input model and the values recovered from the simulated visibilities using the improved TGE for tapering values  $f = 10, 2$  and  $0.6$ , with  $1\sigma$  error bars estimated from  $N_r = 128$  realizations of the simulations. The right panel shows the fractional deviation of the estimated  $C_\ell$  with respect to the input model. Here the shaded region shows the expected statistical fluctuations ( $\sigma_{E_G}/\sqrt{N_r}C_\ell^M$ ) of the fractional deviation for  $f = 0.6$ .

almost 40 per cent. This is possibly a consequence of the fact that the width of the convolution window  $\tilde{w}(\mathbf{U}_g - \mathbf{U}_i)$  increases as the value of  $f$  is reduced, and the variation of the signal amplitude within the width of  $\tilde{w}(\mathbf{U}_g - \mathbf{U}_i)$  becomes important at small baselines, where it is reflected as an overestimate of the value of  $C_\ell$ . Theoretically, we expect the fractional deviation to have random, statistical fluctuations of the order  $\sigma_{E_G}/\sqrt{N_r}C_\ell^M$ , where  $\sigma_{E_G}$  is the standard deviation of the estimated angular power spectrum. The statistical fluctuation expected for  $f = 0.6$  is shown as the shaded region in the right panel of Fig. 2. It can be seen that the fractional deviation is roughly consistent with statistical fluctuations for  $\ell \geq 500$ .

## 2.4 Variance

In the preceding discussion we used several statistically independent realizations of the signal to determine the variance of the estimated binned angular power spectrum. Such a procedure is, by and large, only possible with simulated data. We usually have access to only one statistically independent realization of the input signal, and the aim is to use this not only to estimate the angular power spectrum but also to estimate the uncertainty in the estimated angular power spectrum. In this subsection we present theoretical predictions for the variance of the binned TGE (equation (19)),

$$\sigma_{E_G}^2(a) = \langle \hat{E}_G^2(a) \rangle - \langle \hat{E}_G(a) \rangle^2 \quad (24)$$

which can be used to estimate the uncertainty in the measured angular power spectrum. Following Paper I, we ignore the term  $\sum_i |\tilde{w}(\mathbf{U}_g - \mathbf{U}_i)|^2 |\mathcal{V}_i|^2$  in equation (17) when calculating the variance. The signal contribution from this term to the estimator at the grid point  $\mathbf{U}_g$  scales as  $N_g$ , which is the number of visibilities that contribute to  $\hat{E}_g$ . In comparison to this, the contribution from the term  $|\mathcal{V}_{cg}|^2$  scales as  $N_g^2$ , which is much larger when  $N_g \gg 1$ . Assuming that this condition is satisfied at every grid point that contributes to the binned TGE, it is justifiable to drop the term  $\sum_i |\tilde{w}(\mathbf{U}_g - \mathbf{U}_i)|^2 |\mathcal{V}_i|^2$  when calculating the variance. We then have

$$\sigma_{E_G}^2(a) = \frac{\sum_{gg'} w_g w_{g'} M_g^{-1} M_{g'}^{-1} |\langle \mathcal{V}_{cg} \mathcal{V}_{cg'}^* \rangle|^2}{[\sum_g w_g]^2} \quad (25)$$

which is identical to equation (41) of Paper I, except that we now have the normalization constant  $M_g^{-1}$  instead of  $K_{1g}^{-2}/V_1$ .

It is necessary to model the correlation between the convolved visibilities at two different grid points  $\langle \mathcal{V}_{cg} \mathcal{V}_{cg'}^* \rangle$  in equation (25) in order to make further progress. This correlation is a sum of two parts:

$$\langle \mathcal{V}_{cg} \mathcal{V}_{cg'}^* \rangle = \langle \mathcal{S}_{cg} \mathcal{S}_{cg'}^* \rangle + \langle \mathcal{N}_{cg} \mathcal{N}_{cg'}^* \rangle, \quad (26)$$

the signal and the noise correlation respectively.

Earlier studies (Paper I) showed that we expect the signal correlation  $\langle \mathcal{S}_{cg} \mathcal{S}_{cg'}^* \rangle$  to fall off as  $\exp(-|\Delta \mathbf{U}_{gg'}|^2/\sigma_1^2)$  if the grid separation is increased, where  $\sigma_1 = f^{-1}\sqrt{1+f^2}\sigma_0$  with  $\sigma_0 = 0.76/\theta_{\text{FWHM}}$ . We use this to approximate the signal correlation as

$$\langle \mathcal{S}_{cg} \mathcal{S}_{cg'}^* \rangle = \sqrt{M_g M_{g'}} e^{-|\Delta \mathbf{U}_{gg'}|^2/\sigma_1^2} \tilde{C}_{\tilde{\ell}_a}, \quad (27)$$

where  $\tilde{C}_{\tilde{\ell}_a}$  refers to the angular power spectrum measured at the particular bin  $a$  for which the variance  $\sigma_{E_G}^2(a)$  is being calculated.

The noise correlation

$$\langle \mathcal{N}_{cg} \mathcal{N}_{cg'}^* \rangle = \sum_i \tilde{w}(\mathbf{U}_g - \mathbf{U}_i) \tilde{w}^*(\mathbf{U}_{g'} - \mathbf{U}_i) \langle |\mathcal{N}_i|^2 \rangle \quad (28)$$

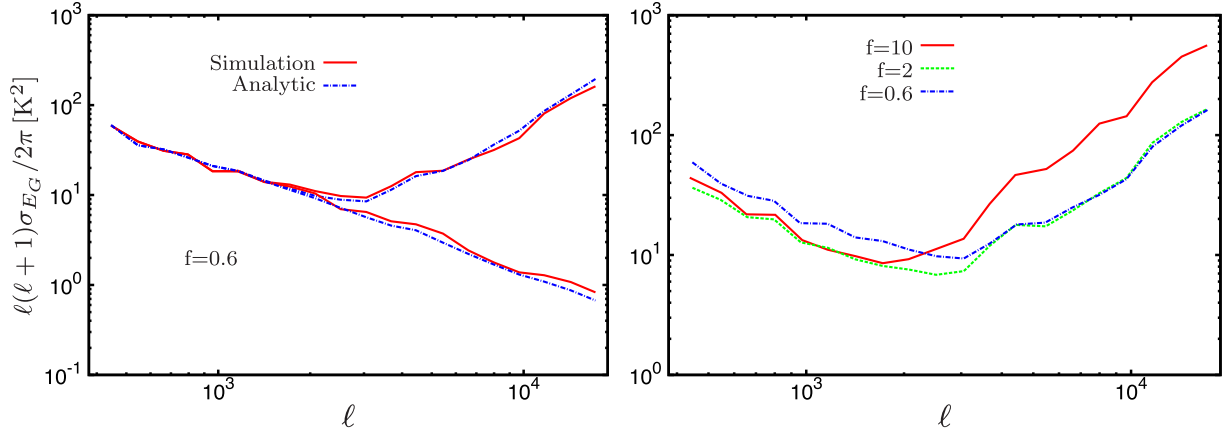
also is expected to decrease as the grid separation is increased, and we model this  $|\Delta \mathbf{U}_{gg'}|$  dependence as

$$\langle \mathcal{N}_{cg} \mathcal{N}_{cg'}^* \rangle = \sqrt{K_{2gg} K_{2g'g'}} e^{-|\Delta \mathbf{U}_{gg'}|^2/\sigma_2^2} (2\sigma_n^2), \quad (29)$$

where  $K_{2gg} = \sum_i |\tilde{w}(\mathbf{U}_g - \mathbf{U}_i)|^2$ ,  $\sigma_2 = 3\sigma_0 f^{-1}$ , and  $\sigma_n^2$  is the variance of the real (and also imaginary) part of  $\mathcal{N}_i$ .

We used equations (29), (27) and (26) in equation (25) to calculate  $\sigma_{E_G}^2(a)$ , the analytic prediction for the variance of the estimated binned angular power spectrum  $\tilde{C}_{\tilde{\ell}_a}$ .

The left panel of Fig. 3 shows the analytic prediction for the variance calculated using equation (25) for a fixed value of  $f = 0.6$ . For comparison, we also show the variance estimated from  $N_r = 128$  independent realizations of the simulated visibilities. We considered two situations: in the first the simulated visibilities only have the signal corresponding to the input model (equation 23) and no system noise, and in the second the visibilities have in addition to the signal a system noise contribution with  $\sigma_n = 1.03$  Jy, which corresponds to a 16-s integration time and a channel width of 125 kHz. It can be seen that the variance calculated from the simulations is dominated by cosmic variance at small  $\ell$  ( $\leq 2,000$ ), and that the variance does not change irrespective of whether we include the system noise or not. The variance calculated from the simulations is dominated



**Figure 3.** In the left panel, the analytic prediction for the variance (equation 25) is compared with the variance estimated from  $N_r = 128$  realizations of the simulated visibilities. Results are shown both with (upper curves) and without (lower curves) the system noise contribution. Both match at small  $\ell$ , where cosmic variance dominates. The system noise, however, is important at large  $\ell$ , where the two sets of results are different. The right panel shows how the variance with system noise obtained from simulations varies for different values of  $f$ .

by the system noise at large  $\ell$  ( $\geq 5000$ ). It can be seen that the analytic predictions are in reasonably good agreement with the values obtained from the simulations over the entire  $\ell$  range that we have considered here. We also considered situations with  $f = 2.0$  and  $10$ , for which comparisons with the analytic results are not shown here. In all cases we found that analytic predictions are in reasonably good agreement with the values obtained from the simulations.

The right panel of Fig. 3 shows how the variance obtained from the simulations changes with  $f$ . It can be seen that at low  $\ell$  the variance increases if the value of  $f$  is reduced. This is a consequence of the fact that cosmic variance increases as the sky response is tapered by reducing  $f$ . The same effect was discussed in detail in our earlier paper (Paper I). It can also be seen that at large  $\ell$  the variance is considerably higher for  $f = 10$  in comparison with  $f = 2$  and  $0.6$ . This  $\ell$  range is dominated by the system noise contribution. The number of independent visibilities that are combined to estimate the power spectrum at any grid point increases as  $f$  is reduced, and this is reflected in a smaller variance as  $f$  is reduced.

### 3 3D $P(k_\perp, k_\parallel)$ ESTIMATION

#### 3.1 3D TGE

We now turn our attention to the redshifted 21-cm H I brightness temperature fluctuations. In this case it is necessary to consider different frequency channels, so equation (1) is generalized to

$$\mathcal{V}_i(v_a) = S(U_i, v_a) + \mathcal{N}_i(v_a). \quad (30)$$

Proceeding in exactly the same manner as for a single frequency channel (equation 2), we have

$$S(U_i, v_a) = \left( \frac{\partial B}{\partial T} \right)_{v_a} \int d^2\theta e^{2\pi i U_i \cdot \theta} \mathcal{A}(\theta, v_a) \delta T(\theta, v_a), \quad (31)$$

and the noises in the different visibility measurements at different frequency channels are uncorrelated:

$$\langle \mathcal{N}_i(v_a) \mathcal{N}_j(v_b) \rangle = \langle |\mathcal{N}_i(v_a)|^2 \rangle \delta_{i,j} \delta_{a,b}. \quad (32)$$

Note that the baseline corresponding to a fixed antenna separation  $U_i = \mathbf{d}_i/\lambda$ , the antenna beam pattern  $\mathcal{A}(\theta, v_a)$  and the factor  $(\partial B/\partial T)_{v_a}$  all vary with the frequency  $v_a$  in equation (31). However,

for the present analysis we consider only the frequency dependence of the H I signal  $\delta T(\theta, v_a)$ , which is assumed to vary much more rapidly with  $v_a$  than the other terms, which are expected to have a relatively slower frequency dependence that has been ignored here. We then have

$$S(U_i, v_a) = \left( \frac{\partial B}{\partial T} \right) \int d^2U \tilde{a}(U_i - U) \Delta \tilde{T}(U, v_a), \quad (33)$$

which is similar to equation (3) introduced earlier.

In equation (33), we can express  $\Delta \tilde{T}(U, v)$  in terms of  $\Delta T(\mathbf{k})$ , which refers to the 3D Fourier decomposition of the H I brightness temperature fluctuations in the region of space from which the redshifted 21-cm radiation originated. We use equation (7) of Bharadwaj & Sethi [2001; or equivalently equation (12) of Bharadwaj & Ali (2005)] to express  $S(U_i, v)$  in terms of the 3D brightness temperature fluctuations:

$$S(U_i, v) = \left( \frac{\partial B}{\partial T} \right) \int \frac{d^3k}{(2\pi)^3} \tilde{a}\left(\mathbf{U}_i - \frac{\mathbf{k}_\perp r}{2\pi}\right) e^{-ik_\parallel r'v} \Delta \tilde{T}(\mathbf{k}), \quad (34)$$

where  $(\mathbf{k}_\perp, k_\parallel)$  are the components of the comoving wavevector  $\mathbf{k}$  respectively perpendicular and parallel to the line of sight,  $r$  is the comoving distance corresponding to the redshifted 21-cm radiation at the observing frequency  $v$ ,  $r' = |dr/dv|$ , and

$$\langle \Delta \tilde{T}(\mathbf{k}) \Delta \tilde{T}^*(\mathbf{k}') \rangle = (2\pi)^3 \delta_D^3(\mathbf{k} - \mathbf{k}') P(\mathbf{k}_\perp, k_\parallel) \quad (35)$$

defines  $P(\mathbf{k}_\perp, k_\parallel)$ , the 3D power spectrum of H I brightness temperature fluctuations.  $v$  here is measured with respect to the central frequency of the observation, and  $r$  and  $r'$  are held fixed at the values corresponding to the central frequency.

We next consider observations with  $N_c$  discrete frequency channels  $v_a$ , with  $a = 0, 1, 2, \dots, N_c - 1$ , each channel of width  $\Delta v_c$ , and the total spanning a frequency bandwidth  $B_{bw}$ . This corresponds to a comoving spatial extent of  $(r'B_{bw})$  along the line of sight, and  $k_\parallel$  now assumes discrete values

$$k_\parallel = \frac{2\pi\tau_m}{r'}, \quad (36)$$

where  $\tau_m$  is the delay variable (Morales & Hewitt 2004; McQuinn et al. 2006), which takes values  $\tau_m = m/B_{bw}$  with  $-N_c/2 < m \leq N_c/2$ . The  $k_\parallel$  integral in equation (34) is now

replaced by a discrete sum  $\int k_{\parallel}/(2\pi) \rightarrow (r' B_{\text{bw}})^{-1} \sum_m$ . It is further convenient to use

$$\mathbf{k}_{\perp} = \frac{2\pi\mathbf{U}}{r}, \quad (37)$$

whereby

$$\begin{aligned} \mathcal{S}(\mathbf{U}_i, v_a) &= \left( \frac{\partial B}{\partial T} \right) \int d^2U \tilde{a}(\mathbf{U}_i - \mathbf{U}) \\ &\times \sum_m e^{-2\pi i \tau_m v_a} \frac{\Delta \tilde{T}(\mathbf{U}, \tau_m)}{B_{\text{bw}} r^2 r'}. \end{aligned} \quad (38)$$

Note here that we can identify  $\tau_m$  as being the Fourier conjugate of  $v_a$ .

We now consider the Fourier transform along the frequency axis of the measured visibilities, which gives the visibilities  $v_i(\tau_m)$  in delay space:

$$v_i(\tau_m) = (\Delta v_c) \sum_a e^{2\pi i \tau_m v_a} \mathcal{V}_i(v_a). \quad (39)$$

The following analysis in this section is based entirely on the delay space visibilities  $v_i(\tau_m)$  defined in equation (39).

Calculating  $s(\mathbf{U}_i, \tau_m)$ , the H I signal contribution to  $v_i(\tau_m)$  using equation (38), we have

$$s(\mathbf{U}_i, \tau_m) = \left( \frac{\partial B}{\partial T} \right) \int d^2U \tilde{a}(\mathbf{U}_i - \mathbf{U}) \left[ \frac{\Delta \tilde{T}(\mathbf{U}, \tau_m)}{r^2 r'} \right], \quad (40)$$

and, rewriting equation (35) in terms of the new variables  $\mathbf{U}$  and  $\tau_m$ , we have

$$\begin{aligned} \langle \Delta \tilde{T}(\mathbf{U}, \tau_m) \Delta \tilde{T}^*(\mathbf{U}, \tau_n) \rangle &= \delta_D^2(\mathbf{U} - \mathbf{U}') \\ &\times [\delta_{m,n} (B_{\text{bw}} r^2 r') P(\mathbf{k}_{\perp}, k_{\parallel})]. \end{aligned} \quad (41)$$

It can be seen that the signals at two different delay channels are uncorrelated. It is also straightforward to verify that the noise contributions  $n_i(\tau_m)$  at two different delay channels are uncorrelated.

In summary of the calculations discussed so far in this section, we see that the visibilities  $v_i(\tau_m)$  at two different delay channels are uncorrelated. It therefore suffices to analyse each delay channel separately, and in the subsequent discussion we restrict our attention to a fixed delay channel  $\tau_m$ . Calculating the correlation of a visibility with itself, we have

$$\begin{aligned} \langle |v_i(\tau_m)|^2 \rangle &= V_0 \left[ \frac{B_{\text{bw}}}{r^2 r'} P(\mathbf{k}_{\perp}, k_{\parallel}) \right] \\ &+ (\Delta v_c)^2 \sum_a \langle |\mathcal{N}_i(v_a)|^2 \rangle. \end{aligned} \quad (42)$$

It is important to note that equations (40), (41) and (42), which hold for a fixed delay channel, are exactly analogous to equations (3), (4) and (7), which hold for a fixed frequency channel. We define the convolved visibilities in exact analogy with equation (9):

$$v_{cg}(\tau_m) = \sum_i \tilde{w}(\mathbf{U}_g - \mathbf{U}_i) v_i(\tau_m), \quad (43)$$

and we define the 3D TGE in exact analogy with equation (17):

$$\begin{aligned} \hat{P}_g(\tau_m) &= \left( \frac{M_g B_{\text{bw}}}{r^2 r'} \right)^{-1} \\ &\times \left( |v_{cg}(\tau_m)|^2 - \sum_i |\tilde{w}(\mathbf{U}_g - \mathbf{U}_i)|^2 |v_i(\tau_m)|^2 \right). \end{aligned} \quad (44)$$

The 3D TGE is, by construction, an unbiased estimator of the 3D power spectrum  $P(\mathbf{k}_{\perp}, k_{\parallel})$ , and we have

$$\langle \hat{P}_g(\tau_m) \rangle = P(\mathbf{k}_{\perp g}, k_{\parallel m}), \quad (45)$$

where  $k_{\parallel m}$  and  $\mathbf{k}_{\perp g}$  are related to  $\tau_m$  and  $\mathbf{U}_g$  through equations (36) and (37), respectively.

### 3.2 Frequency window function

The discrete Fourier transform used to calculate  $v_i(\tau_m)$  in equation (39) assumes that the measured visibilities  $\mathcal{V}_i(v_a)$  are periodic across the frequency bandwidth  $B_{\text{bw}}$  (i.e.  $\mathcal{V}_i(v_a) = \mathcal{V}_i(v_a + B_{\text{bw}})$ ). In reality, the measured visibilities are not periodic over the observational bandwidth, and the discrete Fourier transform encounters a discontinuity at the edge of the band. It is possible to avoid this problem by multiplying the measured visibilities by a frequency window function  $F(v_a)$  that decreases smoothly to zero at the edges of the band. This effectively makes the product  $F(v_a) \mathcal{V}_i(v_a)$  periodic, thereby removing the discontinuity at the edges of the band. This issue was studied by Vedantham et al. (2012) and Thyagarajan et al. (2013), who proposed the Blackman–Nuttall (Nuttall 1981) window function

$$\begin{aligned} F(a) &= c_0 - c_1 \cos\left(\frac{2\pi a}{N_c - 1}\right) + c_2 \cos\left(\frac{4\pi a}{N_c - 1}\right) \\ &- c_3 \cos\left(\frac{6\pi a}{N_c - 1}\right), \end{aligned} \quad (46)$$

where  $c_0 = 0.3635819$ ,  $c_1 = 0.4891775$ ,  $c_2 = 0.1365995$  and  $c_3 = 0.0106411$ . In a recent paper, Chapman et al. (2016) compared different frequency window functions and concluded that the extended Blackman–Nuttall window is the best choice for recovering the H I power spectrum. For the present work, we use the Blackman–Nuttall window as given by equation (46) above. The left panel of Fig. 4 shows the frequency window function for 256 frequency channels spanning a frequency bandwidth of  $B_{\text{bw}} = 16$  MHz, which corresponds to the values that we used in our simulations (discussed later).

We now have

$$v_i^f(\tau_m) = (\Delta v_c) \sum_a e^{2\pi i \tau_m v_a} F(v_a) \mathcal{V}_i(v_a), \quad (47)$$

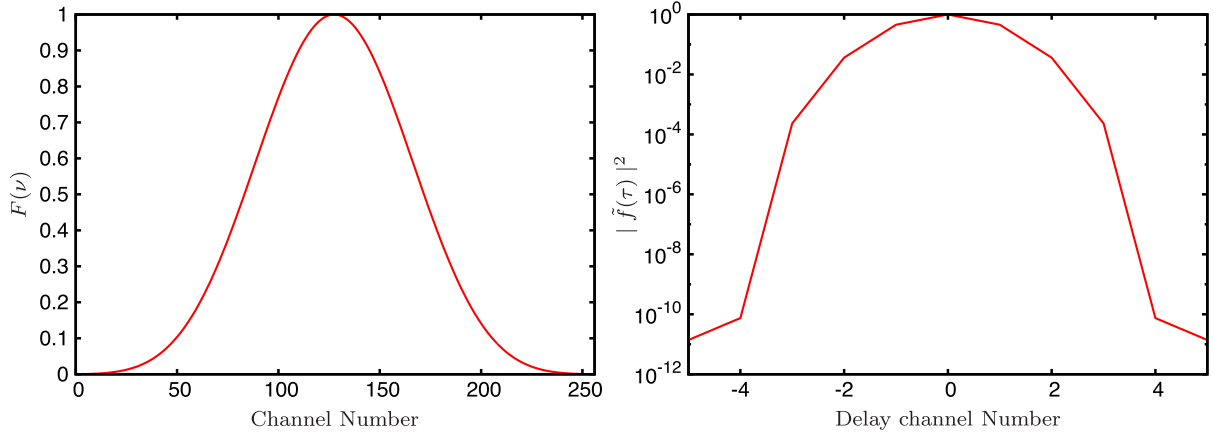
where the  $v_i^f(\tau_m)$  refer to the delay-space visibilities after introducing the frequency window function. The filtered delay-space visibilities  $v_i^f(\tau_m)$  are related to the original delay-space visibilities  $v_i(\tau_m)$  (equation 39) through the convolution

$$v_i^f(\tau_m) = \frac{1}{B_{\text{bw}}} \sum_n \tilde{f}(\tau_m - \tau_n) v_i(\tau_n), \quad (48)$$

where  $\tilde{f}(\tau)$  is the Fourier transform of the frequency window  $F(v)$ . Recall that the delay-space visibilities  $v_i(\tau_m)$  at the different  $\tau_m$  are all independent and uncorrelated. It can be seen, however, that this does not hold for the filtered delay-space visibilities  $v_i^f(\tau_m)$ , for which the different  $\tau_m$  values are correlated, with the extent of this correlation being determined by the width of the function  $\tilde{f}(\tau_m - \tau_n)$  in equation (48). We now use this to calculate the correlation of  $v_i^f(\tau_m)$  at two different values of  $\tau_m$ , for which we have

$$\begin{aligned} \langle v_i^f(\tau_m) v_i^{f*}(\tau_n) \rangle &= \frac{1}{B_{\text{bw}}^2} \sum_a \tilde{f}(\tau_m - \tau_a) \\ &\times \tilde{f}^*(\tau_n - \tau_a) \langle |v_i(\tau_a)|^2 \rangle. \end{aligned} \quad (49)$$





**Figure 4.** The Blackman–Nuttall frequency window  $F(\nu)$  as a function of channel number is shown in the left panel. The right panel shows  $|\tilde{f}(\tau)|^2$ , which is the square of the Fourier transform of  $F(\nu)$ . This is normalized to unity at the central delay channel.

This gives the self-correlation as

$$\langle |v_i^f(\tau_m)|^2 \rangle = \frac{1}{B_{\text{bw}}^2} \sum_a |\tilde{f}(\tau_m - \tau_a)|^2 \langle |v_i(\tau_a)|^2 \rangle. \quad (50)$$

The right panel of Fig. 4 shows  $|\tilde{f}(\tau_m)|^2$  as a function of the delay channel number  $m$ . It can be seen that  $|\tilde{f}(\tau_m)|^2$  has a very narrow extent in delay space, implying that the visibilities  $v_i^f(\tau_m)$  in only three adjacent delay channels are correlated, and that  $v_i^f(\tau_m)$  are uncorrelated if the delay channel separation is larger than this. This also allows us to approximate  $|\tilde{f}(\tau_m - \tau_n)|^2$  using a Kronecker delta function  $\approx B_{\text{bw}}^2 A_f(0) \delta_{m,n}$ , where  $A_f(0) = \frac{1}{B_{\text{bw}}^2} \sum_n |\tilde{f}(\tau_n)|^2$ . The convolution in equation (50) now gives

$$\langle |v_i^f(\tau_m)|^2 \rangle = A_f(0) \langle |v_i(\tau_m)|^2 \rangle. \quad (51)$$

We now generalize this to calculate the correlation for two different values of  $\tau_m$ , which gives

$$\langle v_i^f(\tau_m) v_i^{f*}(\tau_n) \rangle = A_f(m - n) \langle |v_i(\tau_m)|^2 \rangle, \quad (52)$$

where

$$A_f(m - n) = \frac{1}{B_{\text{bw}}^2} \sum_a \tilde{f}(\tau_m - \tau_a) \tilde{f}^*(\tau_n - \tau_a) \quad (53)$$

and  $A_f(m - n) = A_f^*(n - m)$ . We find that  $A_f(m)$  has significant values only for  $m = 0, 1, 2, 3$ , beyond which the values are small; that is, only the visibilities at the three adjacent delay channels have significant correlations, and the visibilities are uncorrelated beyond this separation. We use the self-correlation (equation 51) to calculate the power spectrum estimator later in this subsection, whereas the general expression for the correlation (equation 52) comes in useful for calculating the variance in a subsequent subsection.

Incorporating the frequency window function in the 3D TGE introduces an additional factor of  $A_f(0)$  into the normalization coefficient in equation (44). We now have the final expression for the 3D TGE as

$$\hat{P}_g(\tau_m) = \left( \frac{M_g B_{\text{bw}} A_f(0)}{r^2 r'} \right)^{-1} \times \left( |v_{cg}^f(\tau_m)|^2 - \sum_i |\tilde{w}(\mathbf{U}_g - \mathbf{U}_i)|^2 |v_i^f(\tau_m)|^2 \right). \quad (54)$$

As mentioned earlier,  $\hat{P}_g(\tau_m)$  gives an estimate of the power spectrum  $P(\mathbf{k}_{\perp g}, k_{\parallel m})$ , where  $k_{\parallel m}$  and  $\mathbf{k}_{\perp g}$  are related to  $\tau_m$  and  $\mathbf{U}_g$  through equations (36) and (37), respectively.

### 3.3 Binning and variance

The estimator  $\hat{P}_g(\tau_m)$  presented in equation (54) provides an estimate of the 3D power spectrum  $P(\mathbf{k}_{\perp g}, k_{\parallel m})$  at an individual grid point  $\mathbf{k} = (\mathbf{k}_{\perp g}, k_{\parallel m})$  in the 3D  $\mathbf{k}$ -space. Usually, one would like to average the estimated power spectrum over a bin in  $\mathbf{k}$ -space in order to increase the signal-to-noise ratio. In this section we discuss the bin-averaged 3D TGE and obtain formulas for theoretically predicting the expected variance.

We introduce the binned 3D TGE, which for the bin labelled  $a$  is defined as

$$\hat{P}_G(a) = \frac{\sum_{gm} w_{gm} \hat{P}_g(\tau_m)}{\sum_{gm} w_{gm}}, \quad (55)$$

where the sum is over all the  $\mathbf{k} = (\mathbf{k}_{\perp g}, k_{\parallel m})$  modes, or equivalently the grid points  $(\mathbf{U}_g, \tau_m)$  included in the particular bin  $a$ , and  $w_{gm}$  is the weight assigned to the contribution from any particular grid point. Earlier in this paper, in the discussion immediately following equation (19), we introduced the weighing scheme  $w_g = 1$  in order to calculate  $C_\ell$ . Here we adopt the same scheme  $w_{gm} = 1$  for estimating the 3D power spectrum.

The expectation value of the binned 3D TGE (equation 55),

$$\langle \hat{P}_G(a) \rangle = \bar{P}(\bar{\mathbf{k}}_{\perp}, \bar{k}_{\parallel})_a, \quad (56)$$

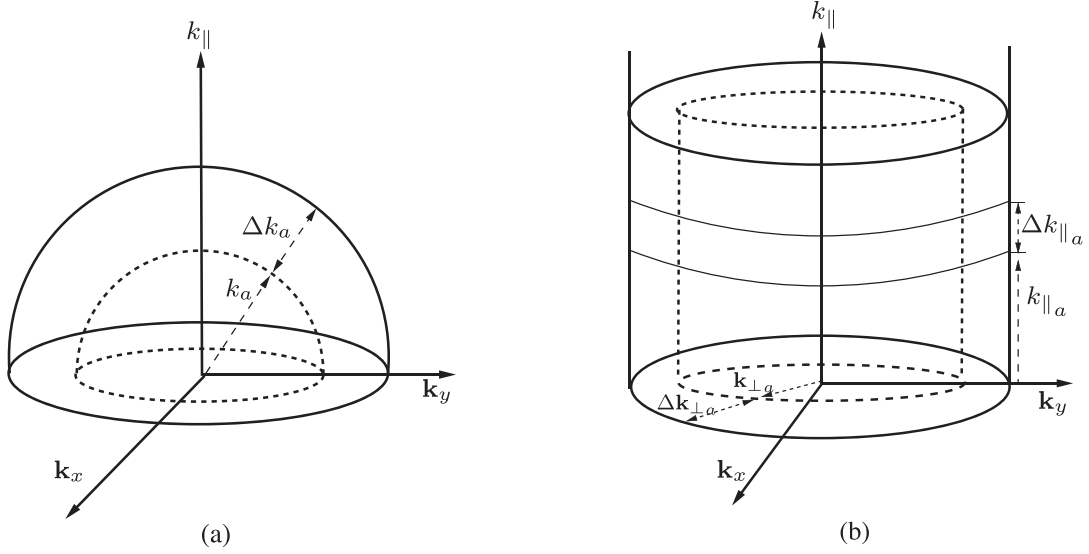
gives an estimate of the bin-averaged 3D power spectrum

$$\bar{P}(\bar{\mathbf{k}}_{\perp}, \bar{k}_{\parallel})_a = \frac{\sum_{gm} w_{gm} P(\mathbf{k}_{\perp g}, k_{\parallel m})}{\sum_{gm} w_{gm}} \quad (57)$$

at

$$(\bar{\mathbf{k}}_{\perp}, \bar{k}_{\parallel})_a = \left( \frac{\sum_{gm} w_{gm} \mathbf{k}_{\perp g}}{\sum_{gm} w_{gm}}, \frac{\sum_{gm} w_{gm} k_{\parallel m}}{\sum_{gm} w_{gm}} \right), \quad (58)$$

where for the particular bin  $a$  the two components  $(\bar{\mathbf{k}}_{\perp}, \bar{k}_{\parallel})_a$  refer to the average wavenumbers respectively perpendicular and parallel to the line of sight. In this paper, we consider two binning schemes, which we discuss later in this subsection. For the moment, we turn our attention to calculating theoretical predictions for the variance of the binned 3D TGE.



**Figure 5.** A typical bin for calculating the spherical power spectrum (left) and the cylindrical power spectrum (right).

The variance calculation closely follows the steps outlined in Section 2.4, and we have the final expression

$$\sigma_{P_G}^2 = \left( \frac{B_{bw} A_f(0)}{r^2 r'} \right)^{-2} \times \frac{\sum_{gm, g'm'} w_{gm} w_{g'm'} M_g^{-1} M_{g'}^{-1} |\langle v_{cg}^f(\tau_m) v_{cg'}^{f*}(\tau_{m'}) \rangle|^2}{[\sum_{gm} w_{gm}]^2}, \quad (59)$$

which closely resembles equation (25), which we used to calculate the variance for  $C_\ell$ , with the difference that we now have a 3D grid instead of the 2D grid utilized earlier for  $C_\ell$ .

It is necessary to model the term  $\langle v_{cg}^f(\tau_m) v_{cg'}^{f*}(\tau_{m'}) \rangle$  in equation (59) to make further progress. The correlation at two different  $\tau_m$  values can be expressed using equation (52) as

$$\langle v_{cg}^f(\tau_m) v_{cg'}^{f*}(\tau_{m'}) \rangle = A_f(m - m') \langle v_{cg}(\tau_m) v_{cg'}^*(\tau_m) \rangle. \quad (60)$$

Following equation (26), we decomposed the correlation  $\langle v_{cg}(\tau_m) v_{cg'}^*(\tau_m) \rangle$  in equation (60) into two parts:

$$\langle v_{cg}(\tau_m) v_{cg'}^*(\tau_m) \rangle = \langle s_{cg}(\tau_m) s_{cg'}^*(\tau_m) \rangle + \langle n_{cg}(\tau_m) n_{cg'}^*(\tau_m) \rangle, \quad (61)$$

corresponding to the signal and to the noise respectively.

We modelled the signal correlation in exact analogy with equation (27) as

$$\langle s_{cg}(\tau_m) s_{cg'}^*(\tau_m) \rangle = \left( \frac{B_{bw}}{r^2 r'} \right) \sqrt{M_g M_{g'}} e^{-|\Delta U_{gg'}|^2 / \sigma_1^2} \bar{P}(\bar{k}_\perp, \bar{k}_\parallel)_a, \quad (62)$$

and the noise correlation is similarly modelled in exact analogy with equation (29) as

$$\langle n_{cg}(\tau_m) n_{cg'}^*(\tau_m) \rangle = (\Delta v_c) B_{bw} \sqrt{K_{2gg} K_{2g'g'}} e^{-|\Delta U_{gg'}|^2 / \sigma_2^2} (2\sigma_n^2). \quad (63)$$

We used equations (63), (62), (61), (60) and (59) to calculate the variance of the binned 3D TGE. In the subsequent analysis we consider two different binning schemes, which we now present.

### 3.3.1 1D spherical power spectrum

The bins here are spherical shells of thickness  $\Delta k_a$ , as shown in the left panel of Fig. 5. The shell thickness will in general vary from bin to bin. The spherical power spectrum  $\bar{P}(\bar{k}_a)$  is obtained by averaging the power spectrum  $P(\mathbf{k})$  over all the  $\mathbf{k}$  modes that lie within the spherical shell corresponding to bin  $a$  shown in the left panel of Fig. 5. The binning here essentially averages out any anisotropy in the power spectrum, and yields the bin-averaged power spectrum as a function of the 1D bin-averaged wavenumber  $\bar{k}_a$ . While we use equation (55) to calculate the bin-averaged power spectrum  $\bar{P}(\bar{k}_a)$ , we calculate the value of  $\bar{k}_a$  using

$$\bar{k}_a = \frac{\sum_{gm} w_{gm} \sqrt{k_{\perp g}^2 + k_{\parallel g}^2}}{\sum_{gm} w_{gm}}. \quad (64)$$

### 3.3.2 2D cylindrical power spectrum

Each bin here is, as shown in the right panel of Fig. 5, an annulus of width  $\Delta k_{\perp a}$  in the  $\mathbf{k}_\perp \equiv (k_x, k_y)$  plane, and it subtends a thickness  $\Delta k_{\parallel a}$  along the third direction  $k_{||}$ . The values of  $\Delta k_{\perp a}$  and  $\Delta k_{\parallel a}$  will, in general, vary from bin to bin. The bins here correspond to sections of a hollow cylinder, and the resulting bin-averaged power spectrum  $\bar{P}(\bar{k}_\perp, \bar{k}_\parallel)_a$  is referred to as the cylindrical power spectrum, which is defined on a 2D space  $(\bar{k}_\perp, \bar{k}_\parallel)_a$ , where the two components refer to the average wavenumbers respectively perpendicular and parallel to the line of sight. The binning of  $P(\mathbf{k})$  here does not assume that the signal is statistically isotropic in the 3D space (i.e. independent of the direction of  $\mathbf{k}$ ). However, the signal is assumed to be statistically isotropic in the plane of the sky, and the binning in  $\mathbf{k}_\perp$  is identical to the binning that we used earlier for  $C_\ell$ . This distinction between  $k_\perp$  and  $k_{||}$  is useful in quantifying the effect of redshift space distortion (Bharadwaj & Sethi 2001; Bharadwaj, Nath & Sethi 2001; Bharadwaj & Ali 2004; Barkana & Loeb 2005; Mao 2012; Majumdar, Bharadwaj & Choudhury 2013; Jensen et al. 2016) and also in distinguishing the foregrounds from the  $H\text{I}$  signal (Morales & Hewitt 2004). We used equations (55) and (58) to calculate  $\bar{P}(\bar{k}_\perp, \bar{k}_\parallel)_a$  and  $(\bar{k}_\perp, \bar{k}_\parallel)_a$  respectively.

#### 4 SIMULATION

In this section we discuss the simulations that we used to validate the 3D power spectrum estimator (equation 54). We start with an input model 3D power spectrum  $P^M(k)$  of redshifted H I 21-cm brightness temperature fluctuations. The aim here is to test how well the estimator is able to recover the input model. For this purpose, the exact form of the input model power spectrum need not mimic the expected cosmological H I signal, and we have used a simple power law

$$P^M(k) = \left(\frac{k}{k_0}\right)^n, \quad (65)$$

which is arbitrarily normalized to unity at  $k = k_0$  and has a power-law index  $n$ . In our analysis we considered  $n = -3$  and  $-2$ , and set  $k_0 = 1 \text{ Mpc}^{-1}$ . The quantity  $\Delta_k^2 = (2\pi^2)^{-1} k^3 P(k)$  provides an estimate of the mean-square brightness temperature fluctuations expected at different length-scales (or equivalently wavenumbers  $k$ ). It can be seen that for  $n = -3$  there is a constant  $\Delta_k^2 = (2\pi^2)^{-1} \text{ K}^2$  across all length-scales, whereas  $\Delta_k^2 = (2\pi^2)^{-1} (k/1 \text{ Mpc}^{-1}) \text{ K}^2$ , which increases linearly with  $k$  for  $n = -2$ . Note that we have used an isotropic input model, in which the power spectrum does not depend on the direction of  $\mathbf{k}$  (i.e.  $P(\mathbf{k}) \equiv P(k)$ ), and the 1D spherical binning and the 2D cylindrical binning are expected to recover the same results.

The simulations were carried out using a  $N^3$  cubic grid of spacing  $L$  covering a comoving volume  $V$ . We use the model power spectrum (equation 65) to generate the Fourier components of the brightness temperature fluctuations corresponding to this grid:

$$\Delta \tilde{T}(\mathbf{k}) = \sqrt{\frac{V P^M(k)}{2}} [a(\mathbf{k}) + ib(\mathbf{k})], \quad (66)$$

where  $a(\mathbf{k})$  and  $b(\mathbf{k})$  are two real-valued independent Gaussian random variables of unit variance. The Fourier transform of  $\Delta T(\mathbf{k})$  yields a single realization of the brightness temperature fluctuations  $\delta T(\mathbf{x})$  on the simulation grid. These fluctuations are, by construction, a Gaussian random field with power spectrum  $P^M(k)$ . We generate different statistically independent realizations of  $\delta T(\mathbf{x})$  by using different sets of random variables  $a(\mathbf{k})$  and  $b(\mathbf{k})$  in equation (66).

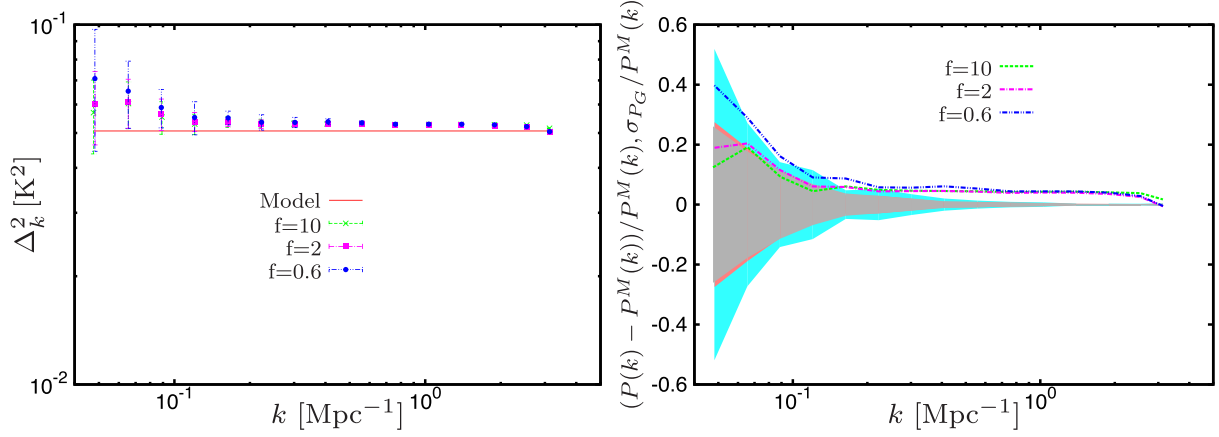
The intention here is to simulate 150-MHz GMRT observations with  $N_c = 256$  frequency channels of width  $(\Delta \nu_c) = 62.5 \text{ kHz}$  covering a bandwidth of  $B_{\text{bw}} = 16 \text{ MHz}$ . This corresponds to H I at redshift  $z = 8.47$  with a comoving distance of  $r = 9.28 \text{ Gpc}$  and  $r' = |dr/d\nu| = 17.16 \text{ Mpc MHz}^{-1}$ . We chose the grid spacing  $L = 1.073 \text{ Mpc}$  so that it exactly matches the channel width  $L = r'_v \times (\Delta \nu_c)$ . We considered a  $N^3 = [2048]^3$  grid, which corresponds to a comoving volume of  $[2197.5 \text{ Mpc}]^3$ . The simulation volume is aligned with the  $z$ -axis along the line of sight, and the two transverse directions are converted to angles relative to the box centre  $(\theta_x, \theta_y) = (x/r, y/r)$ . The transverse extent of the simulation box covers an angular extent that is  $\sim 5$  times the GMRT  $\theta_{\text{FWHM}}$ . The simulation volume corresponds to a frequency width  $\sim 8 \times 16 \text{ MHz}$  along the line of sight. We cut the box into eight equal segments along the line of sight, thus producing eight independent realizations, each subtending  $16 \text{ MHz}$  along the line of sight. The grid index, measured from the distant boundary and increasing towards an observer along the line of sight, was directly converted to the channel number  $\nu_a$  with  $a = 0, 1, 2, \dots, N_c - 1$ . This procedure provides us with  $\delta T(\theta, \nu_a)$ , the brightness temperature fluctuation on the sky in different frequency channels  $\nu_a$ .

We considered 8 h of GMRT observations with a 16-s integration time targeted on an arbitrarily selected field located at RA =  $10^{\text{h}} 46^{\text{m}} 00^{\text{s}}$  and Dec. =  $59^\circ 00' 59''$ . Visibilities were calculated for the simulated baselines corresponding to this observation, for which the  $uv$  coverage is similar to that in fig. 5 of Paper I. The signal contribution to the visibilities  $S(U, \nu_a)$  was calculated by taking the Fourier transform of the product  $(\partial B/\partial T) \times \mathcal{A}(\theta, \nu_a) \times \delta T(\theta, \nu_a)$ , as given by equation (31). The simulations incorporate the fact that the baseline corresponding to a fixed antenna separation  $U_i = \mathbf{d}_i/\lambda$ , the antenna beam pattern  $\mathcal{A}(\theta, \nu_a)$  and the factor  $(\partial B/\partial T)_{\nu_a}$  all vary with the frequency  $\nu_a$  in equation (31). We have  $\sigma_n = 1.45 \text{ Jy}$ , corresponding to a single polarization, with  $\Delta t = 16 \text{ s}$  and  $(\Delta \nu_c) = 62.5 \text{ kHz}$ . However, it is possible to reduce the noise level by averaging an independent data set observed at a different time. Here, we consider a situation in which we average nine independent data sets to reduce the noise level by a factor of 3 to  $\sigma_n = 0.48 \text{ Jy}$ . We carried out the simulations for two cases, (i) no noise ( $\sigma_n = 0 \text{ Jy}$ ), and (ii)  $\sigma_n = 0.48 \text{ Jy}$ . We carried out 16 independent realization of the simulated visibilities in order to estimate the mean power spectrum and its statistical fluctuation (or standard deviation  $\sigma_{P_G}$ ), as presented in the next section.

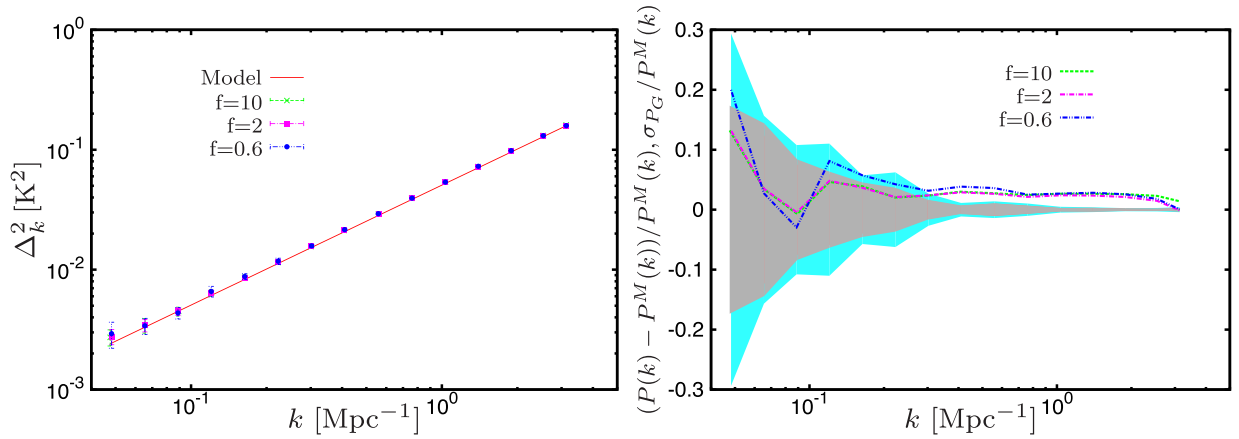
#### 5 RESULTS

The left panels of Figs 6 and 7 show  $\Delta_k^2 = (2\pi^2)^{-1} k^3 P(k)$  for the spherically averaged power spectrum for the power-law index values  $n = -3$  and  $-2$ , respectively. The results are shown for the three values  $f = 10, 2$  and  $0.6$  in order to demonstrate the effect of varying the tapering. The simulations here do not include the system noise contribution. For both  $n = -3$  and  $-2$ , and for all the values of  $f$ , we find that  $\Delta_k^2$  estimated using the 3D TGE is within the  $1 - \sigma_{P_G}$  error bars of the model prediction for the entire range of  $k$  considered here. The right panels of Figs 6 and 7 show the corresponding fractional deviations  $(P(k) - P^M(k))/P^M(k)$ . For comparison, the relative statistical fluctuations,  $\sigma_{P_G}/P^M(k)$ , are also shown by the shaded regions for various values of  $f$ . We find that for both cases,  $n = -3$  and  $-2$ , the fractional deviation is less than 4 per cent at  $k > 0.2 \text{ Mpc}^{-1}$ . The fractional deviation increases as we go to lower- $k$  bins. The fractional deviation also increases if the value of  $f$  is reduced. The maximum fractional deviation has a value of  $\sim 40$  and  $\sim 20$  per cent at the smallest- $k$  bin for  $n = -3$  and  $-2$ , respectively. We find that the fractional deviation is within  $\sigma_{P_G}/P^M(k)$  for  $k \leq 0.3 \text{ Mpc}^{-1}$  and is slightly larger than  $\sigma_{P_G}/P^M(k)$  for  $k \geq 0.3 \text{ Mpc}^{-1}$ . Our results indicate that the 3D TGE is able to recover the model power spectrum to a reasonably good level of accuracy ( $\leq 20$  per cent) at the  $k$  modes  $k \geq 0.1 \text{ Mpc}^{-1}$ . The fractional error at the smaller- $k$  bins increases as the tapering is increased ( $f$  is reduced). Note that a similar behaviour was also found for  $C_\ell$  (Fig. 2). As mentioned earlier, we attribute this discrepancy to the variation of signal amplitude within the width of the convolving window  $\tilde{w}(\mathbf{U}_g - \mathbf{U}_i)$ . This explanation is further substantiated by the fact that the fractional deviation is found to be larger for  $n = -3$ , for which the power spectrum is steeper compared with  $n = -2$ .

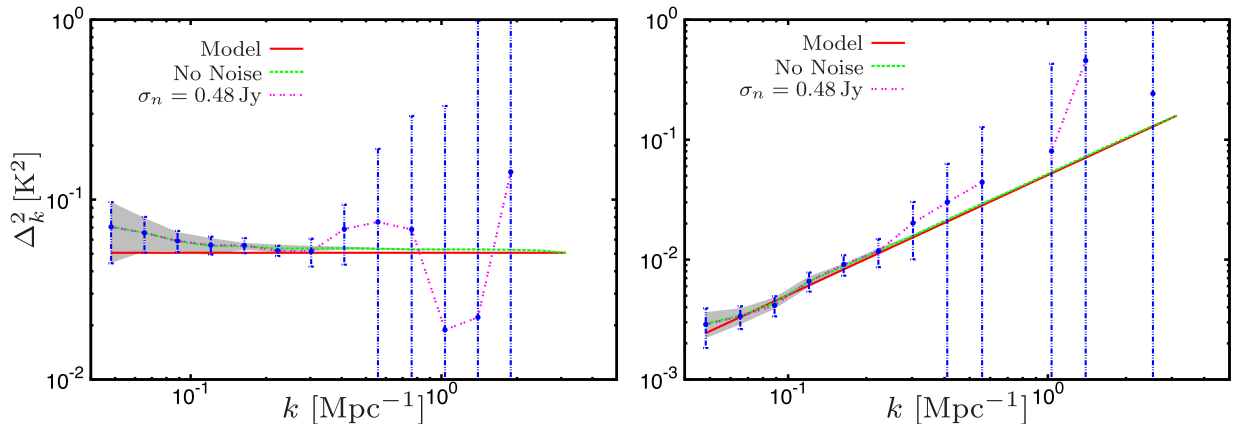
The results thus far have not considered the effect of system noise. We now study how well the 3D TGE is able to recover the input power spectrum in the presence of system noise. The left and right panels of Fig. 8 show the estimated  $\Delta_k^2$  for  $n = -3$  and  $-2$  respectively for the fixed value  $f = 0.6$ . For comparison, we also show the estimated  $\Delta_k^2$  with  $\sigma_n = 0$ . The statistical fluctuations with (without) noise are shown as error bars (shaded regions). It can be



**Figure 6.** The left panel shows the dimensionless power spectrum  $\Delta_k^2$  for various values of  $f$ . The values obtained using the 3D TGE are compared with the model power spectrum for  $n = -3$  and  $\sigma_n = 0$ . The  $1\sigma_{PG}$  error bars were estimated using 16 realizations of the simulated visibilities. The right panel shows the fractional deviation of the estimated power spectrum,  $(P(k) - P^M(k))/P^M(k)$ , relative to the input model  $P^M(k)$  for various values of  $f$ . The relative statistical fluctuations  $\sigma_{PG}/P^M(k)$  are also shown, by the shaded regions.



**Figure 7.** As Fig. 6, but for  $n = -2$ .

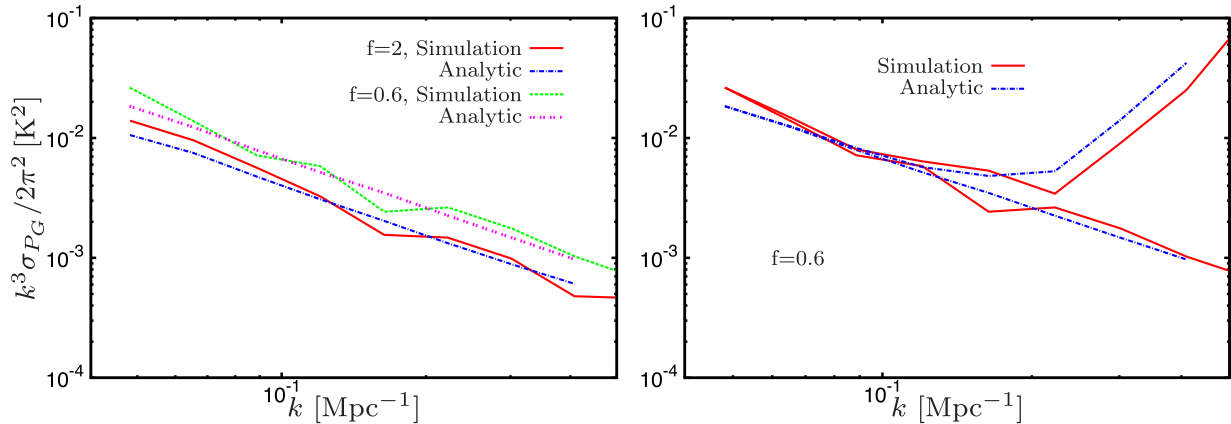


**Figure 8.** The recovered dimensionless power spectrum  $\Delta_k^2$  for  $n = -3$  (left) and  $n = -2$  (right), with and without noise for a fixed value  $f = 0.6$ . The statistical error ( $1\sigma_{PG}$ ) with (without) noise is shown with error bars (shaded region). Note that the estimated  $\Delta_k^2$  has negative values at some of the  $k$ -values in the range for which noise dominates the signal. These data points are not displayed here.

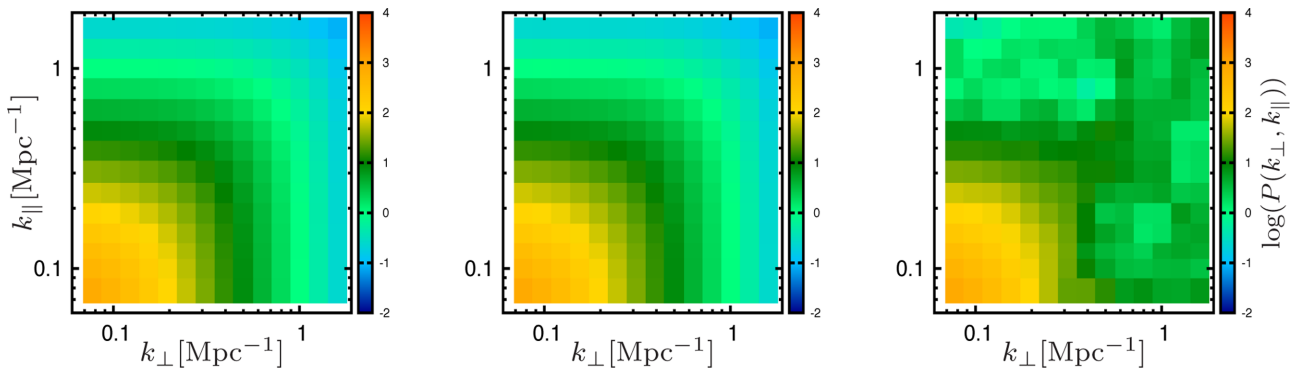
seen that the error is dominated by the cosmic variance at lower values of  $k$  ( $k < 0.2 \text{ Mpc}^{-1}$ ) and that the system noise dominates at larger values of  $k$ . The statistical error exceeds the model power spectrum at large  $k$ , and a statistically significant estimate of the power spectrum is not possible in this range of  $k$ . We are able to

recover the model power spectrum quite accurately at low  $k$ , where  $\sigma_{PG} \leq P^M(k)$ .

We now investigate how well the analytic prediction (equation 59) for  $\sigma_{PG}$  compares with the values obtained from the simulations (Fig. 9) for various values of  $f$ . The number of grid points



**Figure 9.** The left panel shows a comparison of the analytic prediction for the statistical fluctuations of the power spectrum (equation 59) with the simulation for two values of  $f$ ,  $n = -3$  and no system noise. The right panel shows the same comparison with (upper two curves) and without (lower two curves) noise for the fixed value  $f = 0.6$ .



**Figure 10.** The 2D cylindrical power spectrum for  $n = -3$ . The left panel shows the input model power spectrum. The middle and right panels show the estimated power spectrum for  $f = 0.6$  without and with noise respectively.

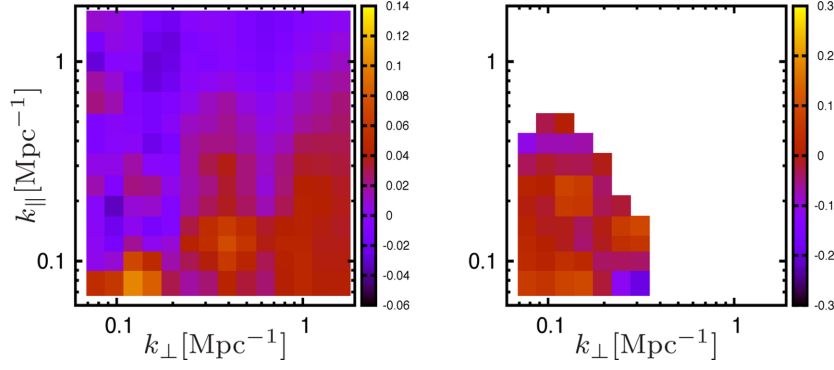
in each  $k$  bin increases with the value of  $k$ , and the computation time also increases with increasing  $k$ . We restricted the  $k$  range to ( $k < 0.4 Mpc^{-1}$ ) in order to keep the computational requirements within manageable limits. The left panel depicts the situation for which there is no system noise. Here, the statistical fluctuations correspond to the cosmic variance. It can be seen that the analytic predictions are in reasonably good agreement with the simulation for both values of  $f$ . We find that the cosmic variance does not change if the value of  $f$  is changed from 2 to 10. As expected, the cosmic variance increases as the sky tapering is increased. The right panel shows the statistical fluctuations with and without noise for the fixed value  $f = 0.6$ . The statistical fluctuations are dominated by the cosmic variance at small values of  $k$  ( $k < 0.2 Mpc^{-1}$ ), and the system noise dominates at large  $k$ . As noted earlier, the statistical fluctuations are well modelled by the analytic predictions in the cosmic variance-dominated regime. We find that our analytic prediction somewhat overestimates  $\sigma_{P_G}$  in the noise-dominated region. This overestimate possibly originates from the noise modelling in equation (59). We plan to investigate this in future work.

Up until now we have discussed the results for the 1D spherical power spectrum; we now present the results for the 2D cylindrical power spectrum. We use 15 equally spaced logarithmic bins in both the  $k_\perp$  and  $k_\parallel$  directions in order to estimate the 2D cylindrical power spectrum. Fig. 10 shows the 2D cylindrical power spectrum  $P(k_\perp, k_\parallel)$  using the 3D TGE. The left panel shows the input model for  $n = -3$ . The middle and right panels respectively show the estimated

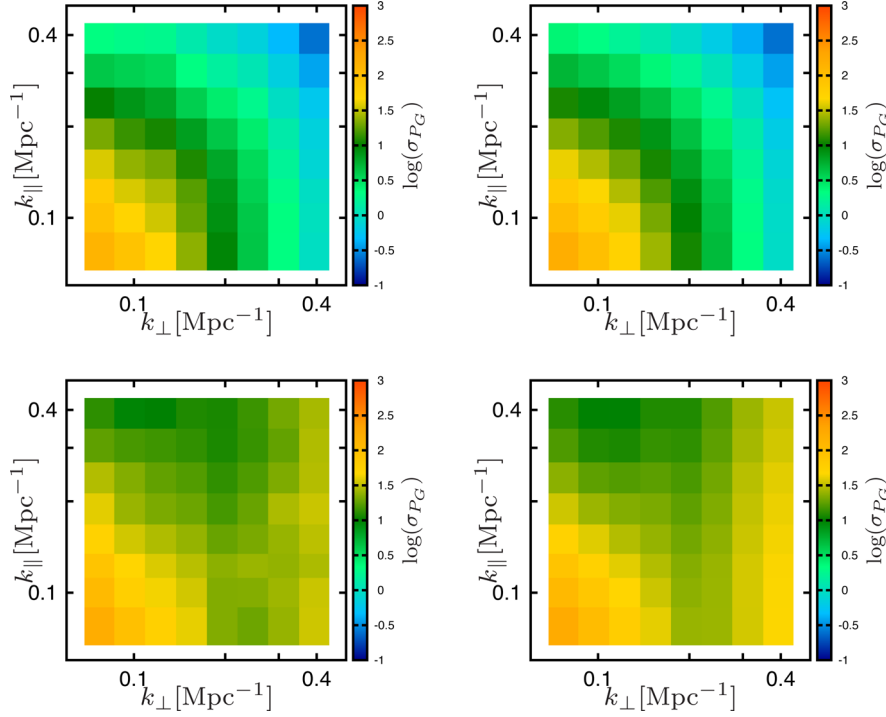
power spectrum with  $f = 0.6$  for situations in which the system noise is not included and included in the simulated visibilities. The left and middle panels appear almost identical, indicating that the 3D TGE is able to recover the input model power spectrum accurately across the entire  $(k_\perp, k_\parallel)$  range. We find that we are able to recover the model power spectrum in the limited range  $k_\perp \lesssim 0.5 Mpc^{-1}$  and  $k_\parallel \lesssim 0.5 Mpc^{-1}$  in the presence of system noise. Fig. 11 shows the fractional deviation  $(P^M(k_\perp, k_\parallel) - P(k_\perp, k_\parallel))/P(k_\perp, k_\parallel)$  for  $f = 0.6$ ; here the left and right panels show the results without and with system noise respectively. From the left panel it can be seen that the fractional deviation is less than 14 per cent for the entire  $k$  range when the system noise is not included in the simulation. We find that it is not possible to recover the power spectrum reliably at large  $k$  when the system noise is included. The right panel shows the fractional deviation only where it is less than 30 per cent. The fractional deviation exceeds 100 per cent at large  $k$ , and these values have not been shown.

We now investigate how well the analytic prediction (equation 59) for  $\sigma_{P_G}$  compares with the values obtained from the simulations (Fig. 12) for  $f = 0.6$ . The two upper panels depict the situation in which there is no system noise, with the left and right panels respectively showing the simulated and the analytic prediction for the statistical fluctuation  $\sigma_{P_G}$ . We find that the analytic predictions match quite well with the simulation for the entire  $k$  range. The two lower panels depict the situation for which the system noise is included, with the left and right panels respectively showing the

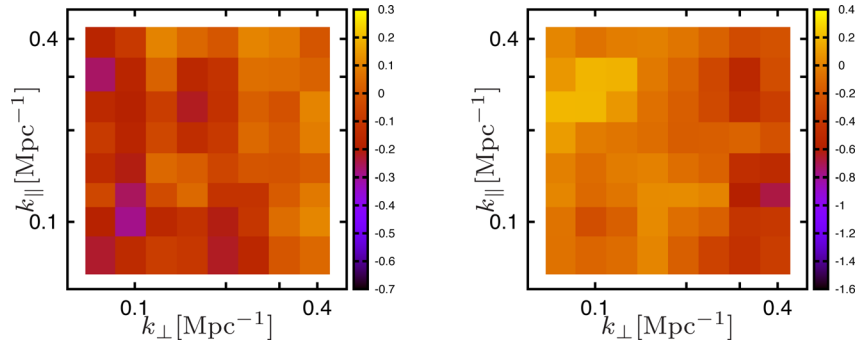




**Figure 11.** The left and right panels show the fractional deviation  $(P^M(k_\perp, k_\parallel) - P(k_\perp, k_\parallel))/P(k_\perp, k_\parallel)$  without and with noise, respectively, for  $n = -3$  and  $f = 0.6$ .



**Figure 12.** The statistical fluctuation ( $\sigma_{PG}$ ) for the 2D cylindrical power spectrum for  $n = -3$  and  $f = 0.6$ . The upper and lower panels show the results without and with system noise respectively, and the left and right panels show the results from the simulations and the analytic prediction respectively.



**Figure 13.** The left and right panels show the fractional deviation of  $\sigma_{PG}$  without and with system noise respectively.

simulated and the analytic prediction for  $\sigma_{PG}$ . The left and right panels of Fig. 13 show the fractional deviation between the simulated and analytic  $\sigma_{PG}$  without and with system noise respectively. We find that we have less than 20 per cent fractional deviation in 73

and 64 per cent of the bins in  $(k_\perp, k_\parallel)$  space without and with system noise respectively. The fractional deviation shows a greater spread in values when the system noise is included, compared with the situation without system noise. We do not, however, find any obvious

pattern in the distribution of the bins that show a high fractional deviation.

## 6 SUMMARY AND CONCLUSIONS

Quantifying the statistical properties of the diffuse sky signal directly from the visibilities measured in low-frequency radio-interferometric observation is an important task. In this paper we have presented a statistical estimator, namely the tapered gridded estimator (TGE), which has been developed for this purpose. The measured visibilities are here gridded in the  $uv$ -plane in order to reduce the complexity of the computation. The contribution from the discrete sources in the periphery of the telescope's FoV, particularly the sidelobes, poses a problem for power spectrum estimation. The TGE suppresses the contribution from the outer regions by tapering the sky response through a suitably chosen window function. The TGE also internally estimates the noise bias from the input data, and subtracts this out to give an unbiased estimate of the power spectrum. In addition to the mathematical formalism for the estimator and its variance, we have also presented simulations of 150-MHz GMRT observations, which are used to validate the estimator.

We first considered a situation in which we have observations at a single frequency for which the 2D TGE provides an estimate of the angular power spectrum  $C_\ell$ . The work here presents an improvement over an earlier version of the 2D TGE presented in [Paper I](#). This is important in the context of the diffuse Galactic synchrotron emission, which is one of the major foregrounds for the cosmological 21-cm signal. Apart from this, the diffuse Galactic synchrotron emission is a probe of the cosmic ray electrons and the magnetic fields in the ISM of our own Galaxy, and this is an important area of research in its own right.

It is necessary also to include the frequency variation of the sky signal in order to quantify the cosmological 21-cm signal. Here the 3D TGE provides an estimate of  $P(k)$ , the power spectrum of the 21-cm brightness temperature fluctuations. We considered two binning schemes, which provide the 1D spherical power spectrum  $P(k)$  and the 2D cylindrical power spectrum  $P(k_\perp, k_\parallel)$  respectively. In all cases, we found that the TGE is able to accurately recover the input model used for the simulations. The analytic predictions for the variance were also found to be in reasonably good agreement with the simulations in most situations.

Foregrounds are possibly the biggest challenge for detecting the cosmological 21-cm power spectrum. Various studies (e.g. [Datta et al. 2010](#)) have shown that the foreground contribution to the cylindrical power spectrum  $P(k_\perp, k_\parallel)$  is expected to be restricted to within a wedge in the  $(k_\perp, k_\parallel)$  plane. The extent of this 'foreground wedge' is determined by the angular extent of the telescope's FoV. In principle, it is possible to limit the extent of the foreground wedge by tapering the telescope's FoV. In the context of estimating the angular power spectrum  $C_\ell$ , our earlier work ([Paper II](#)) demonstrated that the 2D TGE is able to suppress the contribution from the outer parts and the sidelobes of the telescope's beam pattern. We have not explicitly considered the foregrounds in our analysis of the 3D TGE presented in this paper. However, we expect the 3D TGE to suppress the contribution from the outer parts and the sidelobes of the telescope beam pattern while estimating the power spectrum  $P(k_\perp, k_\parallel)$ , thereby reducing the area in the  $(k_\perp, k_\parallel)$ -plane under the foreground wedge.

The 3D TGE holds the promise of allowing us to reduce the extent of the foreground wedge by tapering the sky response. It is, however, necessary to note that this comes at a cost, which we now discuss. First, we lose information at the largest angular scales

owing to the reduced FoV. This restricts the smallest  $k$ -value at which it is possible to estimate the power spectrum. Second, the reduced FoV results in a larger cosmic variance for the smaller angular modes that are within the tapered FoV. The actual value of the tapering parameter  $f$  that would be used to estimate  $P(k_\perp, k_\parallel)$  will possibly be determined by optimizing between the cosmic variance and the foreground contribution. A possible strategy would be to use different values of  $f$  for different bins in the  $(k_\perp, k_\parallel)$ -plane. It should also be noted that the effectiveness of the tapering proposed here depends on the actual baseline distribution, and a reasonably dense  $uv$  coverage is required for a proper implementation of the TGE. We propose to include foregrounds in the simulations and to address these issues in future work. We also plan to apply this estimator to 150-MHz GMRT data in the future.

## ACKNOWLEDGEMENTS

S. Choudhuri would like to acknowledge the University Grants Commission, India for providing financial support through a Senior Research Fellowship. S. Chatterjee is supported by a University Grants Commission Research Fellowship. SSA would like to acknowledge CTS, IIT Kharagpur for the use of its facilities and to thank the authorities of the IUCAA, Pune, India for providing the Visiting Associateship programme.

## REFERENCES

- Ali S. S., Bharadwaj S., 2014, *JA&A*, 35, 157
- Ali S. S., Bharadwaj S., Chengalur J. N., 2008, *MNRAS*, 385, 2166
- Ali Z. S. et al., 2015, *ApJ*, 809, 61
- Bandura K. et al., 2014, in Stepp L. M., Gilmozzi R., Hall H. J. eds, *Proc. SPIE Vol. 9145, Ground-based and Airborne Telescopes V*. SPIE, Bellingham, p. 914522
- Barkana R., Loeb A., 2005, *ApJ*, 624, L65
- Bernardi G. et al., 2009, *A&A*, 500, 965
- Bharadwaj S., Ali S. S., 2004, *MNRAS*, 352, 142
- Bharadwaj S., Ali S. S., 2005, *MNRAS*, 356, 1519
- Bharadwaj S., Sethi S. K., 2001, *JA&A*, 22, 293
- Bharadwaj S., Nath B., Sethi S. K., 2001, *JA&A*, 22, 21
- Bowman J. D., Morales M. F., Hewitt J. N., 2009, *ApJ*, 695, 183
- Bowman J. D. et al., 2013, *Publ. Astron. Soc. Aust.*, 30, e031
- Bull P., Ferreira P. G., Patel P., Santos M. G., 2015, *ApJ*, 803, 21
- Chapman E. et al., 2012, *MNRAS*, 423, 2518
- Chapman E., Zaroubi S., Abdalla F. B., Dulwich F., Jelić V., Mort B., 2016, *MNRAS*, 458, 2928
- Choudhuri S., Bharadwaj S., Ghosh A., Ali S. S., 2014, *MNRAS*, 445, 4351 (Paper I)
- Choudhuri S., Bharadwaj S., Roy N., Ghosh A., Ali S. S., 2016, *MNRAS*, 459, 151 (Paper II)
- Datta A., Bowman J. D., Carilli C. L., 2010, *ApJ*, 724, 526
- Di Matteo T., Perna R., Abel T., Rees M. J., 2002, *ApJ*, 564, 576
- Dillon J. S. et al., 2014, *Phys. Rev. D*, 89, 023002
- Dillon J. S. et al., 2015, *Phys. Rev. D*, 91, 123011
- Furlanetto S. R., Oh S. P., Briggs F. H., 2006, *Phys. Rep.*, 433, 181
- Ghosh A., Bharadwaj S., Ali S. S., Chengalur J. N., 2011a, *MNRAS*, 411, 2426
- Ghosh A., Bharadwaj S., Ali S. S., Chengalur J. N., 2011b, *MNRAS*, 418, 2584
- Ghosh A., Prasad J., Bharadwaj S., Ali S. S., Chengalur J. N., 2012, *MNRAS*, 426, 3295
- Iacobelli M. et al., 2013, *A&A*, 558, A72
- Jacobs D. C. et al., 2016, *ApJ*, 825, 114
- Jelić V. et al., 2008, *MNRAS*, 389, 1319
- Jensen H., Majumdar S., Mellema G., Lidz A., Iliev I. T., Dixon K. L., 2016, *MNRAS*, 456, 66

- Koopmans L. et al., 2015, Proc. Sci., Advancing Astrophysics with the Square Kilometre Array. SISSA, Trieste, PoS (AASKA14), 1
- Liu A., Tegmark M., 2012, MNRAS, 419, 3491
- Liu A., Parsons A. R., Trott C. M., 2014a, Phys. Rev. D, 90, 023018
- Liu A., Parsons A. R., Trott C. M., 2014b, Phys. Rev. D, 90, 023019
- McQuinn M., Zahn O., Zaldarriaga M., Hernquist L., Furlanetto S. R., 2006, ApJ, 653, 815
- Majumdar S., Bharadwaj S., Choudhury T. R., 2013, MNRAS, 434, 1978
- Mao X.-C., 2012, ApJ, 744, 29
- Mellema G. et al., 2013, Exper. Astron., 36, 235
- Morales M. F., Hewitt J., 2004, ApJ, 615, 7
- Morales M. F., Wyithe J. S. B., 2010, ARA&A, 48, 127
- Neben A. R. et al., 2016, ApJ, 826, 199
- Nuttall A. H., 1981, IEEE Trans. Acoustics, Speech & Signal Proc., ASSP-29, 84
- Paciga G. et al., 2011, MNRAS, 413, 1174
- Paciga G. et al., 2013, MNRAS, 433, 639
- Parsons A. R. et al., 2010, AJ, 139, 1468
- Parsons A. R., Pober J. C., Aguirre J. E., Carilli C. L., Jacobs D. C., Moore D. F., 2012, ApJ, 756, 165
- Parsons A. R. et al., 2014, ApJ, 788, 106
- Planck Collaboration XIII, 2016, A&A, 594, A13
- Pober J. C. et al., 2013a, AJ, 145, 65
- Pober J. C. et al., 2013b, ApJ, 768, L36
- Pober J. C. et al., 2014, ApJ, 782, 66
- Pober J. C. et al., 2016, ApJ, 819, 8
- Prasad P., Subrahmanya C. R., 2011, Exper. Astron., 31, 1
- Pritchard J. R., Loeb A., 2012, Rep. Prog. Phys., 75, 086901
- Santos M. G., Cooray A., Knox L., 2005, ApJ, 625, 575
- Shaver P. A., Windhorst R. A., Madau P., de Bruyn A. G., 1999, A&A, 345, 380
- Thyagarajan N. et al., 2013, ApJ, 776, 6
- Thyagarajan N. et al., 2015, ApJ, 807, L28
- Tingay S. et al., 2013, Publ. Astron. Soc. Aust., 30, 7
- Trott C. M., Wayth R. B., Tingay S. J., 2012, ApJ, 757, 101
- Trott C. M. et al., 2016, ApJ, 818, 139
- van Haarlem M. P. et al., 2013, A&A, 556, A2
- Vedantham H., Udaya Shankar N., Subrahmanyan R., 2012, ApJ, 745, 176
- Yatawatta S. et al., 2013, A&A, 550, 136

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.